

From: Ford, Vanessa
S nt: Thursday, October 11, 2001 1:28 PM
To: STIC-Biotech/ChemLib
Subject: In re: 09596101 Sequence Search

Please search SEQ ID NO: 1.
Please search SEQ ID NO: 3.
Please include an interference search.

Vanessa L. Ford
Biotechnology Patent Examiner
Office: CM1 8D17 ✓
Mailbox: CM1 8E12 ✓
Phone: 703.308.4735

Art unit Please

RECEIVED
OCT 11 2001
STIC

P

SEARCH REQUEST FORM

Access DB#

52795

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

For Sequence Searches Only Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

STAFF USE ONLY

Searcher: Al Nasser
Searcher Phone #: 308-4581
Searcher Location: Biotec lib.
Date Searcher Picked Up: 10/11/01
Date Completed: 10/11/01
Searcher Prep & Review Time: _____
Clinal Prep Time: 6 min
Online Time: 2 min

Type of Search

NA Sequence (#) _____
AA Sequence (#) 2
Structure (#) _____
Bibliographic _____
Litigation _____
Fulltext _____
Patent Family _____
Other _____

Vendors and cost where applicable

STN _____
Dialog _____
Questel/Orbit _____
Dr Link _____
Lexis/Nexis _____
Sequence Systems ABSS02
WWW/Internet _____
Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:48:17 ; Search time 35.17 Seconds
(without alignments)
22.409 Million cell updates/sec

Title: US-09-596-101A-1
Perfect score: 64
Sequence: 1 SGEIDITFTGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

al number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601:*

- 1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:*
- 2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:*
- 3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:*
- 4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:*
- 5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:*
- 6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:*
- 7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:*
- 8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:*
- 9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:*
- 10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:*
- 11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:*
- 12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:*
- 13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:*
- 14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:*
- 15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:*
- 16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:*
- 17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:*
- 18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:*
- 19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:*
- 20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:*
- 21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:*
- 22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64	100.0	13	AA24915	Eisenia foetida co
2	64	100.0	20	AA24914	Eisenia foetida co
3	50	78.1	306	AA97362	Oerskovia beta-1.3
4	45	70.3	294	AA952463	Mycobacterium tube
5	44	68.8	263	AA29455	Oerskovia xanthine
6	44	68.8	303	AA29457	Oerskovia xanthine
7	44	68.8	435	AA29456	Oerskovia xanthine
8	39	60.9	261	AA11599	Beta-1.3-glucanase
9	39	60.9	292	AA88406	Trichoderma harzia
10	38	59.4	223	AA23318	Corn beta-carotene
11	37	57.8	275	AA644125	Arabidopsis thalia

12	37	57.8	324	21	AA644124	Arabidopsis thalia
13	37	57.8	336	21	AA644123	Arabidopsis thalia
14	36	56.2	222	22	AAE00115	Bacillus lichenifo
15	36	56.2	419	22	AA870788	N. magadali bacter
16	36	56.2	1149	20	AA97640	Apaf-1XL/Delta2-10
17	36	56.2	1194	20	AA91071	Apoptosis inducer
18	36	56.2	1194	21	AA94943	Human full-length
19	36	56.2	1194	22	AA97649	Wild type Apaf-1 p
20	36	56.2	1205	20	AA91072	Apoptosis inducer
21	36	56.2	1205	22	AA97637	Apaf-1L protein se
22	36	56.2	1237	22	AA97638	Apaf-1L protein s
23	36	56.2	1248	22	AA97636	Apaf-1XL protein s
24	36	56.2	1248	22	AA97639	Apaf-1XL-M368L pro
25	36	56.2	1248	22	AA97641	Apaf-1XL-L10A prot
26	36	56.2	1248	22	AA97642	Apaf-1XL-D27A prot
27	36	56.2	1248	22	AA97643	Apaf-1XL-E39Q prot
28	36	56.2	1248	22	AA97644	Apaf-1XL-K63Q prot
29	36	56.2	1248	22	AA97645	Apaf-1XL-L83A prot
30	36	56.2	1248	22	AA97646	Apaf-1XL-K160R prot
31	36	56.2	1248	22	AA97647	Apaf-1XL-M368L/K16
32	36	54.7	175	19	AA86077	S. pneumoniae deri
33	35	54.7	176	19	AA85821	S. pneumoniae deri
34	35	54.7	408	21	AA836267	Arabidopsis thalia
35	35	54.7	430	18	AA836265	Arabidopsis thalia
36	35	54.7	430	18	AA836267	Arabidopsis thalia
37	35	54.7	430	19	AA836266	Arabidopsis thalia
38	35	54.7	438	21	AA836265	Arabidopsis thalia
39	35	54.7	485	21	AA836265	Arabidopsis thalia
40	35	54.7	653	21	AA836265	Arabidopsis thalia
41	35	54.7	690	21	AA836265	Arabidopsis thalia
42	35	54.7	690	21	AA836265	Arabidopsis thalia
43	35	54.7	718	21	AA836265	Arabidopsis thalia
44	35	54.7	783	21	AA836265	Arabidopsis thalia
45	35	54.7	1196	20	AA836265	Arabidopsis thalia

ALIGNMENTS

RESULT 1	
AA24915	
ID	AA24915 standard; peptide: 13 AA.
XX	AA24915;
AC	
XX	
DT	25-AUG-1999 (first entry)
XX	
DE	Eisenia foetida coelomic cytolytic factor 1 peptide.
XX	
KW	Eisenia foetida; coelomic cytolytic factor 1; CCF-1; cancer;
KW	trypanosomal infection; bacterial infection; tumour therapy;
KW	Inflammation; Immunology.
XX	
OS	Eisenia foetida.
XX	
PN	WO9931229-A2.
XX	
PD	24-JUN-1999.
XX	
PF	16-DEC-1998; 98WO-EP08169.
XX	
PR	17-DEC-1997; 97EP-0203974.
XX	
PA	(VLA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX	
PI	De Baetselier P;
XX	
DR	WPI; 1999-385905/32.
XX	
PT	Eisenia foetida polypeptides derived from coelomic cytolytic factor 1
XX	
PS	Claim 1; Page 45; 49pp; English.
XX	

CC LIG109 is useful for degrading or modifying beta-glucan-contg.
 CC material. Its amino acid sequence was deduced from a genomic
 CC DNA sequence (AAT29043) isolated from an O. xanthineolytica library.
 CC Recombinant beta-1,3-glucanase can be produced on a large scale
 CC using transformed host cells, esp. Bacillus subtilis DN1885 or
 CC TOC46. Protease-free beta-1,3-glucanase can be obtd. that is useful
 CC for lysing fungal cell walls, allowing recovery of intracellular
 CC proteins. The enzyme is also useful for the prepn. of protoplasts
 CC and for the prodn. of pigments, colorants, flavours, yeast extract
 CC and pharmaceuticals.

SO Sequence 306 AA;

Query Match 78.1%; Score 50; DB 17; Length 306;
 Best Local Similarity 75.0%; Pred. No. 0.12;

Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDITETIGN 12
 |||||:|:|
 180 sgeidimenvgn 191

RESULT 4

ID AAB52463 standard; protein; 294 AA.

AC AAB52463;

DT 23-FEB-2001 (first entry)

DE Mycobacterium tuberculosis secreted protein #28.

KW Mycobacterium tuberculosis secreted protein; MTSP; vaccine.

OS Mycobacterium tuberculosis.

PN W0200066143-A1.

PD 09-NOV-2000.

PF 04-MAY-2000; 2000MO-US12197.

PR 04-MAY-1999; 9905-0132479.

PR 04-MAY-1999; 9905-0132503.

PA (PUBL-) PUBLIC HEALTH RES INST NEW YORK.

Gennaro ML, Gomez MJ;

WPI; 2001-007151/01.

Novel Mycobacterium tuberculosis secreted polypeptides and
 polynucleotides useful in diagnosis, treatment and prophylaxis of
 tuberculosis

Claim 11; Fig 1; 60pp; English.

The present invention relates to Mycobacterium tuberculosis secreted
 CC proteins (MTSP), where the polypeptide has M. tuberculosis specific
 CC antigenic and immunogenic properties. Compositions of the invention may
 CC be useful for diagnosing Mycobacterium tuberculosis infection and as a
 CC vaccine against M. tuberculosis infection.

SO Sequence 294 AA;

Query Match 70.3%; Score 45; DB 22; Length 294;
 Best Local Similarity 75.0%; Pred. No. 1;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGEIDITETIGN 12
 |||||:|:|

Db 169 sgeidimenvgn 180

RESULT 5

ID AAM29455 standard; Protein; 263 AA.

AC AAM29455;

DT 14-APR-1998 (first entry)

DE Oerskovia xanthineolytica mature beta-1,3-glucanase.

KW Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
 fungal cell wall; intracellular product; purification; protoplast.

OS Oerskovia xanthineolytica LIG109 (DSM 10297).

Key Location/Qualifiers

FT CDS 23..955

FT sig_peptide 23..120

FT mat_peptide 164..952

FT /*tag= c

W09739114-A1.

PD 23-OCT-1997.

PF 14-APR-1997; 97WO-DK00160.

PR 23-AUG-1996; 96DK-0000885.

PR 12-APR-1996; 96DK-0000427.

PA (NOVO) NOVO-NORDISK AS.

PI Diers I, Ferrer P, Halkier T, Hedegaard L;

WPI; 1997-526451/48.

DR N-PSDB; NAT89155.

Example 2; Page 35-36; 64pp; English.

This polypeptide comprises a novel Oerskovia xanthineolytica (OX)
 CC enzyme that exhibits beta-1,3-glucanase (BG) activity. Its amino
 CC acid sequence was deduced from an isolated genomic DNA sequence
 CC (see AAT89155). Claimed DNA constructs that encode the novel BG (see
 CC also AAM29456 for corrected sequence), a mannose binding domain (see
 CC AAM29458) or a full-length enzyme, i.e. BG with mannose binding
 CC domain (see AAM29456), can be used to produce recombinant BG
 CC polypeptides, with or without a mannose binding domain, in fungal
 CC or bacterial host cells. BG polypeptides are used for the
 CC degradation or modification of beta-glucan containing material,
 CC especially for the gentle lysis of microbial cell walls, thereby
 CC enabling recovery of desirable intracellular products with a
 CC reduced amount of contaminants. They can also be used for the
 CC production of e.g. pigments, colourants, flavourants, yeast
 CC extracts, pharmaceuticals, food or feed compositions, and to
 CC prepare protoplasts for use in fusion, transformation and cloning
 CC studies.

SO Sequence 263 AA;

Query Match 68.8%; Score 44; DB 18; Length 263;
 Best Local Similarity 72.7%; Pred. No. 1.4;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 8; SGEIDITETIGN 12
 |||||:|:|

OY 1 SGEIDITETIG 11
 |||||:|:|
 Db 116 sgeidmetenvg 126

RESULT 6

AAW29457
 ID AAW29457 standard; Protein; 303 AA.

AC AAW29457;

DT 14-APR-1998 (first entry)

DE Oerskovia xanthineolytica beta-1,3-glucanase.

KW Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
 fungal cell wall; intracellular product; purification; protoplast.

OS Oerskovia xanthineolytica LG109 (DSM 10297).

Key Location/Qualifiers

FT Peptide 1..52
 FT /label= Sig_peptide
 FT 53..303
 FT Protein /label= Mat_protein

XX MO9739114-A1.

XX 23-OCT-1997.

XX 14-APR-1997; 97WO-DK00160.

XX 23-AUG-1996; 96DK-0000885.

XX 12-APR-1996; 96DK-0000427.

XX (NOVO) NOVO-NORDISK AS.

XX Diers I, Ferrer P, Halkier T, Hedegard L;

XX MPI; 1997-526451/48.

XX N-PSDB; AAT89157.

PT New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
 xanthineolytica, used particularly for the lysis of microbial cells
 for obtaining desirable products

PS Example 2; Page 42-43; 64pp; English.

CC This sequence comprises the polypeptide precursor of a novel
 Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
 (BG) activity. Its amino acid sequence was deduced from an
 isolated genomic DNA sequence (see AAT89157). Claimed DNA constructs
 CC that encode the novel BG (see also AAW29457), a mannose binding
 domain (see AAW29458) or a full-length enzyme, i.e. BG with mannose
 CC binding domain (see AAW29456), can be used to produce recombinant BG
 CC polypeptides, with or without a mannose binding domain, in fungal
 CC or bacterial host cells. BG polypeptides are used for the
 CC degradation or modification of beta-glucan containing material,
 CC especially for the gentle lysis of microbial cell walls, thereby
 CC enabling recovery of desirable intracellular products with a
 CC reduced amount of contaminants. They can also be used for the
 CC production of e.g. pigments, colourants, flavourants, yeast
 CC extracts, pharmaceuticals, food or feed compositions, and to
 CC prepare protoplasts for use in fusion, transformation and cloning
 CC studies.

CC Sequence 303 AA;

Query Match 68.8%; Score 44; DB 18; Length 303;
 Best Local Similarity 72.7%; Pred. No. 1.7;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDITETIG 11
 |||||:|:|
 Db 169 sgeidmetenvg 179

RESULT 7

AAW29456
 ID AAW29456 standard; Protein; 435 AA.

AC AAW29456;

DT 14-APR-1998 (first entry)

DE Oerskovia xanthineolytica beta-1,3-glucanase.

KW Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
 fungal cell wall; intracellular product; purification; protoplast.

OS Oerskovia xanthineolytica LG109 (DSM 10297).

Key Location/Qualifiers

FT Peptide 1..52
 FT /label= Sig_peptide
 FT 53..435
 FT Protein /label= Mat_protein

XX Domain 304..435
 /label= Mannose-binding_domain

XX MO9739114-A1.

XX 23-OCT-1997.

XX 14-APR-1997; 97WO-DK00160.

XX 23-AUG-1996; 96DK-0000885.

XX 12-APR-1996; 96DK-0000427.

XX (NOVO) NOVO-NORDISK AS.

XX Diers I, Ferrer P, Halkier T, Hedegard L;

XX MPI; 1997-526451/48.

XX N-PSDB; AAT89156.

PT New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia
 xanthineolytica, used particularly for the lysis of microbial cells
 for obtaining desirable products

PS Example 2; Page 39-40; 64pp; English.

CC This sequence comprises the polypeptide precursor of a novel
 Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase
 (BG) activity and which includes a mannose binding domain (MBD).
 CC Its amino acid sequence was deduced from an isolated genomic DNA
 CC sequence (see AAT89156). Claimed DNA constructs that encode the
 CC novel BG lacking a MBD (see AAW29455 and AAW29457), a MBD (see
 CC AAW29458), or the full-length enzyme can be used to produce recombinant
 CC BG polypeptides, with or without a mannose binding domain, in fungal
 CC or bacterial host cells. BG polypeptides are used for the
 CC degradation or modification of beta-glucan containing material,
 CC especially for the gentle lysis of microbial cell walls, thereby
 CC enabling recovery of desirable intracellular products with a
 CC reduced amount of contaminants. They can also be used for the
 CC production of e.g. pigments, colourants, flavourants, yeast
 CC extracts, pharmaceuticals, food or feed compositions, and to
 CC prepare protoplasts for use in fusion, transformation and cloning
 CC studies.

CC Sequence 435 AA;

Query Match 68.8%; Score 44; DB 18; Length 435;
 Best Local Similarity 72.7%; Pred. No. 2.5;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDIETIG 11
 |||||:|:|
 Db 169 sgeidimenvy 179

RESULT 8

AAR1599
 ID AAR1599 standard; Protein: 261 AA.
 AC AAR1599;

DT 18-JUN-1991 (first entry)

DE Beta-1,3-glucanase.

KM Alkalophilic; heat resistant.

XX Bacillus sp.

Q5 JP03053883-A.

XX PD 07-MAR-1991.

XX PF 20-JUL-1989; 89JP-0185928.

XX PR 20-JUL-1989; 89JP-0185928.

XX PA (SHKJ) SHINGIUTSU KAIHATSU.

XX DR WPI; 1991-113290/16.

XX DR N-PSDB; AAQ11293.

PT Heat resistant beta-1,3-glucanase gene DNA - derived from

XX alkali-compatible Bacillus sp.

PS Claim 7; Fig 2; 11pp; Japanese.

XX The new enzyme is derived from alkalophilic Bacillus sp. The

CC enzyme acts on beta-1,3-glucans to form glucose and laminaribiose.

CC It is stable over a wide pH range with opt. activity at weakly

CC acidic pH and has good heat resistance.

XX Sequence 261 AA;

Query Match 60.9%; Score 39; DB 12; Length 261;
 Best Local Similarity 63.6%; Pred. No. 13;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIG 12
 |||||:|:|
 Db 119 geidimenvy 129

RESULT 9

AAR88406
 ID AAR88406 standard; Protein: 292 AA.

XX AAR88406;

DT 19-AUG-1996 (first entry)

DE Trichoderma harzianum endo-1,3(4)-beta-glucanase.

XX Yeast; fungus; endo-beta-glucanase; EC-3.2.1.6; hydrolase;

KM endo-1,3(4)-beta-glucanase; beta-glucan degradation; hydrolysis;

XX detergent; surfactant; fungicide; antifungal; cleaning.

XX OS Trichoderma harzianum CBS 243.71.

XX PN W09531533-A1.

XX PD 23-NOV-1995.
 XX 11-MAY-1995; 95WO-DK00188.
 XX 11-MAY-1994; 94DK-0000546.

XX (NOVO) NOVO-NORDISK AS.

XX Andersen LN, Breinholt J, Christgau S, Dalboge H;

PI Kauppinen MS, Kofoed LV, Olsen HS;

XX WPI; 1996-010920/01.

DR N-PSDB; AAT09876.

XX DNA encoding endo-beta-glucanase from Trichoderma harzianum -

PT useful, e.g., in food processing, as antifungal agent, in cleaning

PT compns., etc.

XX Claim 1; Page 40; 57pp; English.

XX This polypeptide is immunologically reactive with antibodies

CC raised against a pure endo-1,3(4)-beta-glucanase from T. harzianum

CC CBS 243.71. Using this sequence, the protein may be expressed

CC recombinantly in transformed host cells, particularly Aspergillus

CC oryzae or Aspergillus niger. Typical applications of the

CC recombinantly produced protein are preparation of protoplasts

CC and yeast extracts, brewing, wine and press-juice manufacture, in

XX foods and feeds, as antifungal agents, etc.

XX Sequence 292 AA;

Query Match 60.9%; Score 39; DB 17; Length 292;
 Best Local Similarity 77.8%; Pred. No. 14;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 GEIDIETIG 10
 |||||:|:|
 Db 158 geidimenvy 166

RESULT 10

AAV32318
 ID AAV32318 standard; Protein: 223 AA.

XX AAV32318;

DT 28-FEB-2000 (first entry)

DE Corn beta-carotene hydroxylase.

XX Beta-carotene hydroxylase; corn; maize; transgenic plant;

KM carotenoid; zeaxanthin; pigment.

XX Zea mays;

XX WO955887-A2.

XX WO955887-A2.

XX 04-NOV-1999.

XX 16-APR-1999; 99WO-US08384.

XX 24-APR-1998; 98US-0083042.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Cahoon RE, Kinney AJ, Pearlstein RW, Williams ME;

XX WPI; 2000-062037/05.

XX N-PSDB; AA234967.

XX Novel carotenoid biosynthesis enzyme polynucleotides and polypeptides


```

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159339.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

```

```

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match: 57.88; Score 37; DB 21; Length 275;
Best Local Similarity 50.08; Pred. No. 32;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 3 EIDIIETIGN 12
   |||: |||
Db 75 evdYIqvIghn 84

RESULT 12
AA04124
ID AA04124 standard; Protein; 324 AA.
XX AC
XX AA04124;
XX AC
XX 18-OCT-2000 (first entry)
XX DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 55230.
XX KW
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS
XX Arabidopsis thaliana.
XX PN
XX EP1033405-A2.
XX PD
XX 06-SEP-2000.
XX XX
XX 25-FEB-2000; 2000EP-0301439.
XX PF
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 29-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.
XX 01-APR-1999; 99US-0127462.
XX 06-APR-1999; 99US-0128234.
XX 08-APR-1999; 99US-0128714.
XX 16-APR-1999; 99US-0129845.
XX 19-APR-1999; 99US-0130077.
XX 21-APR-1999; 99US-0130449.
XX 23-APR-1999; 99US-0130510.
XX 23-APR-1999; 99US-0130891.
XX 28-APR-1999; 99US-0131449.
XX 30-APR-1999; 99US-0132048.
XX 30-APR-1999; 99US-0132407.
XX 04-MAY-1999; 99US-0132484.
XX 05-MAY-1999; 99US-0132485.
XX 06-MAY-1999; 99US-0132486.

```

PR	06-MAY-1999;	9905-0132487.
PR	07-MAY-1999;	9905-0132863.
PR	11-MAY-1999;	9905-0134256.
PR	14-MAY-1999;	9905-0134218.
PR	14-MAY-1999;	9905-0134219.
PR	14-MAY-1999;	9905-0134321.
PR	14-MAY-1999;	9905-0134370.
PR	18-MAY-1999;	9905-0134376.
PR	19-MAY-1999;	9905-0134941.
PR	20-MAY-1999;	9905-0135124.
PR	21-MAY-1999;	9905-0135353.
PR	24-MAY-1999;	9905-0135629.
PR	25-MAY-1999;	9905-0136021.
PR	27-MAY-1999;	9905-0136392.
PR	28-MAY-1999;	9905-0136782.
PR	01-JUN-1999;	9905-0137222.
PR	03-JUN-1999;	9905-0137528.
PR	04-JUN-1999;	9905-0137502.
PR	07-JUN-1999;	9905-0137724.
PR	08-JUN-1999;	9905-0138094.
PR	10-JUN-1999;	9905-0138540.
PR	10-JUN-1999;	9905-0138847.
PR	14-JUN-1999;	9905-0139119.
PR	16-JUN-1999;	9905-0139452.
PR	17-JUN-1999;	9905-0139453.
PR	17-JUN-1999;	9905-0139499.
PR	18-JUN-1999;	9905-0139454.
PR	18-JUN-1999;	9905-0139455.
PR	18-JUN-1999;	9905-0139456.
PR	18-JUN-1999;	9905-0139457.
PR	18-JUN-1999;	9905-0139458.
PR	18-JUN-1999;	9905-0139459.
PR	18-JUN-1999;	9905-0139460.
PR	18-JUN-1999;	9905-0139461.
PR	18-JUN-1999;	9905-0139462.
PR	18-JUN-1999;	9905-0139463.
PR	18-JUN-1999;	9905-0139750.
PR	21-JUN-1999;	9905-0139817.
PR	22-JUN-1999;	9905-0139899.
PR	23-JUN-1999;	9905-0140353.
PR	23-JUN-1999;	9905-0140354.
PR	24-JUN-1999;	9905-0140699.
PR	28-JUN-1999;	9905-0140823.
PR	29-JUN-1999;	9905-0140991.
PR	30-JUN-1999;	9905-0141287.
PR	01-JUL-1999;	9905-0141842.
PR	01-JUL-1999;	9905-0142154.
PR	02-JUL-1999;	9905-0142055.
PR	06-JUL-1999;	9905-0142390.
PR	08-JUL-1999;	9905-0142803.
PR	09-JUL-1999;	9905-0142920.
PR	12-JUL-1999;	9905-0142977.
PR	13-JUL-1999;	9905-0143542.
PR	14-JUL-1999;	9905-0143624.
PR	15-JUL-1999;	9905-0144003.
PR	16-JUL-1999;	9905-0144085.
PR	16-JUL-1999;	9905-0144086.
PR	19-JUL-1999;	9905-0144325.
PR	19-JUL-1999;	9905-0144331.
PR	19-JUL-1999;	9905-0144332.
PR	19-JUL-1999;	9905-0144333.
PR	19-JUL-1999;	9905-0144334.
PR	19-JUL-1999;	9905-0144335.
PR	20-JUL-1999;	9905-0144352.
PR	20-JUL-1999;	9905-0144632.
PR	21-JUL-1999;	9905-0144884.
PR	21-JUL-1999;	9905-0144814.
PR	21-JUL-1999;	9905-0145086.
PR	22-JUL-1999;	9905-0145085.
PR	22-JUL-1999;	9905-0145087.
PR	22-JUL-1999;	9905-0145089.
PR	22-JUL-1999;	9905-0145192.
PR	23-JUL-1999;	9905-0145145.
PR	23-JUL-1999;	9905-0145218.
PR	23-JUL-1999;	9905-0145224.
PR	26-JUL-1999;	9905-0145276.
PR	27-JUL-1999;	9905-0145913.
PR	27-JUL-1999;	9905-0145918.
PR	27-JUL-1999;	9905-0145919.
PR	28-JUL-1999;	9905-0145951.
PR	02-AUG-1999;	9905-0146386.
PR	02-AUG-1999;	9905-0146388.
PR	02-AUG-1999;	9905-0146389.
PR	03-AUG-1999;	9905-0147038.
PR	04-AUG-1999;	9905-0147204.
PR	04-AUG-1999;	9905-0147302.
PR	05-AUG-1999;	9905-0147192.
PR	05-AUG-1999;	9905-0147260.
PR	06-AUG-1999;	9905-0147303.
PR	06-AUG-1999;	9905-0147416.
PR	09-AUG-1999;	9905-0147493.
PR	09-AUG-1999;	9905-0147993.
PR	10-AUG-1999;	9905-0148171.
PR	11-AUG-1999;	9905-0148319.
PR	12-AUG-1999;	9905-0148341.
PR	13-AUG-1999;	9905-0148565.
PR	13-AUG-1999;	9905-0148684.
PR	16-AUG-1999;	9905-0149368.
PR	17-AUG-1999;	9905-0149173.
PR	18-AUG-1999;	9905-014942


```

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```

```

PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match : 57.88; Score 37; DB 21; Length 336;
Best Local Similarity 50.08; Pred. No. 41;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 3 EIDIERIGN 12
Db 136 evdvidvign 145

RESULT 14
AAE00115
ID AAE00115 standard; Protein; 222 AA.
XX
AC AAE00115;
XX
DT 31-MAY-2001 (first entry)
XX
DE Bacillus licheniformis RP-II protease variant (T109R).
XX
KW Residual protease II; RP-II; additive; cleaning composition; detergent;
mutant; muten; variant.
XX
OS Bacillus licheniformis.
XX
FH Key Location/Qualifiers
FT Msc-difference 109
FT /note= "Wild type Thr substituted with Arg"
XX
PN W0200116285-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-DK00476.
XX
PR 31-AUG-1999; 99DK-0001212.
PR 20-OCT-1999; 99DK-0001500.
XX
PA (NOVO ) NOVOZYMES AS.
XX
PI Norregaard-madsen M, Rahbek Ostergaard P, Voge Christensen CB;
PI Plensted lassen S;
XX
DR WPI; 2001-226680/23.
XX
PT Novel RP-II type protease and its variants useful as constituents in
XX detergent compositions, additives and cleaning compositions
XX
PS Example 3; Page -; 132pp; English.
XX
CC The present sequence is Bacillus licheniformis (BLC) RP-II protease
CC variant (T109R).
CC The patent discloses RP-II (Residual protease II) protein, their
CC corresponding nucleic acid sequences and variants. RP-II protease
CC is useful as a constituent in additives, detergent compositions and
CC other cleaning compositions, optionally in combination with other
CC enzymes such as proteases, lipases, cellulases, amylases, peroxidases
CC or oxidases. The variants of RP-II have improved properties such as
CC substrate specificities, catalytic rate, stability, especially towards
CC the action of proteolytic enzymes and improved resistance towards
CC peroxidase.
CC Note: This sequence is not shown in the specification but is derived
CC from the mature B. licheniformis RPII protein [SEQ ID NO: 2] shown in
CC page 101-102 of the specification (AAE00011). The specification also
CC refers to following variants: (a) V1F, (b) D7G+T125+E152G+N182I,

```

CC (c) S119A+S144T+O157R+S175I, (d) H141A, (e) E152A+O122R, (f) N182T+C185A,
 CC (g) E152A+V144L and (h) E104K+O204R. However the positions of Val in (a)
 CC and (g), Asn residue in (b) and (f), Ser residues in (c), His residue in
 CC (d), Gln residue in (e) and (h) do not match with the B. licheniformis
 CC native wild type RPII protease sequence.

SO Sequence 222 AA;

Query Match 56.2%; Score 36; DB 22; Length 222;
 Best Local Similarity 58.3%; Pred. No. 39;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 2 GEIDIIETIGNR 13
 | | | | |
 Db 98 galelseplgnr 109

RESULT 15

AAB70788
 AAB70788 standard; Protein; 419 AA.

AC AAB70788;

BT 25-MAY-2001 (first entry)

DE N. magadali bacteriophage phi-CH1 6-methyltransferase.

Genome; halophilic; polyhydroxybutyrate; inducible expression;
 6-methyltransferase.

KW Bacteriophage phi-CH1.
 OS Natrilaiba magadali.

XX DE19937719-A1.

XX 22-FEB-2001.

XX 10-AUG-1999; 99DE-1037719.

XX 10-AUG-1999; 99DE-1037719.

XX (LDB1/) LDBRITZ W.

XX Witte A, Baranyi U, Klein R;

XX WPI; 2001-245930/26.

XX N-PSDB; AAF61284.

DR New nucleic acid from phage phiCH1, used to create vectors for
 expressing proteins and polymers in halophilic Archaea -

PS Claim 9; Page 41-42; 72pp; German.

This invention describes a novel isolated nucleic acid (I) from the
 genome of phage phi-CH1 of Natrilaiba magadali. The genome of phi-CH1 is a
 combination of 48300 and 10198 base pair sequences (S1 and S2), both
 fully defined in the specification. The invention also describes a
 novel (1) recombinant vector (RV1) containing at least one copy of (I);
 (2) recombinant vector (RV2) containing the ori of phi-CH1; (3) cell
 transformed with RV1 or RV2; (4) isolated polypeptide (II) encoded by
 (I); N. magadali cells free from the prophage of phi-CH1; (5) phage
 variants having lytic properties different from those of wild-type
 phi-CH1 and having genomic sequences at least 70 % homologous with the
 phi-CH1 genome; (6) use of phi-CH1 as gene transfer vector; and (7) use
 of halophilic Archaea for production of proteins and other polymers.
 Vectors that contain (I) are used to transform halophilic Archaea,
 specifically N. magadali, for production of proteins and other polymers
 (e.g. poly(hydroxybutyrate)). Vectors containing (I) allow controlled,
 inducible expression of compounds in Archaea.

SO Sequence 419 AA;

Query Match 56.2%; Score 36; DB 22; Length 419;
 Best Local Similarity 54.5%; Pred. No. 81;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
 QY 3 EIDIIETIGNR 13
 | | | | |
 Db 202 didissetlgnr 212

Search completed: October 11, 2001, 15:51:38
 Job time: 201 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: October 11, 2001, 15:48:37 ; Search time 19.14 seconds
(without alignments)
13.985 Million cell updates/sec

Title: US-09-596-101A-1

Sequence: 1 SEIDIIITIGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 197339 segs, 20590346 residues
Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PCPNUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	78.1	306	2	US-08-824-707-2
2	39	60.9	226	2	US-08-712-072C-4
3	39	60.9	292	2	US-08-737-526-4
4	39	60.9	292	4	US-09-098-580-4
5	37	57.8	262	1	US-08-392-828C-37
6	37	57.8	262	3	US-09-330-945-37
7	37	57.8	321	2	US-08-712-072C-3
8	37	54.7	430	1	US-08-476-008-44
9	35	54.7	430	1	US-08-306-063-44
10	35	54.7	430	1	US-08-833-485-44
11	35	54.7	430	4	US-09-137-440-44
12	34	53.1	654	1	US-08-392-828C-2
13	34	53.1	654	3	US-09-330-945-2
14	34	53.1	1119	4	US-09-396-651B-2
15	34	53.1	2409	6	5180808-2
16	33	51.6	394	1	US-07-621-193A-5
17	33	51.6	394	1	US-08-018-489C-5
18	33	51.6	434	2	US-09-008-962-1
19	33	51.6	434	2	US-08-675-507-1
20	33	51.6	434	3	US-09-213-205-1
21	33	51.6	761	1	US-07-906-395-2
22	33	51.6	761	1	US-08-192-632-2
23	33	51.6	761	1	US-08-710-676-2
24	33	51.6	761	1	US-09-099-902B-2
25	33	51.6	761	5	PCT-US93-06080-2
26	33	51.6	887	1	US-07-867-106-3
27	33	51.6	4551	3	US-09-330-878-1

28	32	50.0	77	2	US-08-343-443B-13	Sequence 13, Appl
29	32	50.0	114	1	US-08-031-399-3	Sequence 3, Appl
30	32	50.0	114	1	US-08-393-305-6	Sequence 6, Appl
31	32	50.0	114	1	US-08-726-817-6	Sequence 6, Appl
32	32	50.0	114	1	US-08-504-042-3	Sequence 3, Appl
33	32	50.0	114	2	US-08-725-969-6	Sequence 6, Appl
34	32	50.0	114	2	US-08-794-524-6	Sequence 6, Appl
35	32	50.0	114	4	US-09-189-193-6	Sequence 6, Appl
36	32	50.0	114	5	PCT-US94-03793-3	Sequence 3, Appl
37	32	50.0	122	1	US-08-300-903A-3	Sequence 3, Appl
38	32	50.0	162	1	US-08-031-399-2	Sequence 2, Appl
39	32	50.0	162	1	US-08-393-305-5	Sequence 5, Appl
40	32	50.0	162	1	US-08-284-393B-9	Sequence 9, Appl
41	32	50.0	162	1	US-08-726-817-5	Sequence 5, Appl
42	32	50.0	162	1	US-08-504-042-2	Sequence 2, Appl
43	32	50.0	162	2	US-08-725-969-5	Sequence 5, Appl
44	32	50.0	162	2	US-08-794-524-5	Sequence 5, Appl
45	32	50.0	162	4	US-09-189-193-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-824-707-2
Sequence 2, Application US/08824707
Patent No. 5919688
GENERAL INFORMATION:
APPLICANT: Ferrer, Pau
APPLICANT: Diets, Ivan
APPLICANT: Hedegaard, Lisbeth
APPLICANT: Halkier, Torben
APPLICANT: Asenjo, Juan
APPLICANT: Savva, Demetris
TITLE OF INVENTION: No. 5919688e1 enzyme with beta-1,3-glucanase activity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: No. 5919688o No. 5919688dsk of No. 5919688ch America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,707
FILING DATE: 14-April-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4290,204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-824-707-2

Query Match: 78.1%; Score 50; DB 2; Length 306;
Best Local Similarity 75.0%; Pred. No. 0.09;
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 SEIDIIITIGN 12

APPLICATION NUMBER: 08/737,526
FILING DATE: 08-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Valeta, Gregg A
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4174.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 292 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-098-580-4

Query Match 60.8%; Score 39; DB 4; Length 292;
Best Local Similarity 77.8%; Pred. No. 9.3;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 GSEIDIIETI 10
DB 158 GSEIDIMETV 166

RESULT 5
US-08-392-828C-37
Sequence 37, Application US/08392828C
Patent No. 5795962
GENERAL INFORMATION:
APPLICANT: IMANAGA, SADAKI
APPLICANT: MUTA, TATSUSHI
APPLICANT: SEKI, NORIKI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
ADDRESS: THIBEAULT
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,828C
FILING DATE: 28-FEB-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein

LOCATION: 1..262
OTHER INFORMATION: /note="BGI A1 SEQUENCE (FIGURE 2)."
US-08-392-828C-37

Query Match 57.8%; Score 37; DB 1; Length 262;
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GSEIDIIETIG 11
DB 130 GSEIDVMEARG 140

RESULT 6
US-09-330-945-37
Sequence 37, Application US/09330945
Patent No. 6077946
GENERAL INFORMATION:
APPLICANT: IMANAGA, SADAKI
APPLICANT: MUTA, TATSUSHI
APPLICANT: SEKI, NORIKI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB
TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
ADDRESS: THIBEAULT, LLP
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/330,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/119,995
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: FJN-032DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 262 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..262
OTHER INFORMATION: /note="BGI A1 SEQUENCE (FIGURE 2)."
US-09-330-945-37

Query Match 57.8%; Score 37; DB 3; Length 262;
Best Local Similarity 63.6%; Pred. No. 19;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GSEIDIIETIG 11
DB 158 GSEIDIMETV 166

Db 130 SGEIDVMEARG 140

RESULT 7

US-08-712-072C-3

; Sequence 3, Application US/08712072C
; Patent No. 5925541

; GENERAL INFORMATION:

; APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng

; TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amster, Rothstein & Eberstein

; STREET: 90 Park Avenue

; CITY: New York

; STATE: NY

; COUNTRY: US

; ZIP: 10016

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/712,072C

; FILING DATE: 11-SEP-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Bogosian, Elizabeth A.

; REGISTRATION NUMBER: 39,911

; REFERENCE/DOCKET NUMBER: 63475/97

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 697-5995

; TELEFAX: (212) 286-0854 or 286-0082

; TELEX: TWX 710-581-4766

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 321 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE:

; ORIGINAL SOURCE:

; ORGANISM: el3b, Bacillus circulans

; 08-712-072C-3

QY 1 SGEIDIETIG 11

Db 189 SGEIDVMEARG 199

RESULT 8

US-08-476-008-44

; Sequence 44, Application US/08476008
; Patent No. 5627061

; GENERAL INFORMATION:

; APPLICANT: Barry, Gerard F.

; APPLICANT: Kishore, Ganesh M.

; APPLICANT: Padgett, Stephen R.

; APPLICANT: Stallings, William C.

; TITLE OF INVENTION: Glycosyltransferase-3-Phosphate Synthases

; 08-476-008-44

Query Match 57.8%; Score 37; DB 2; Length 321;

Best Local Similarity 63.6%; Pred. No. 24;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,008

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/306,063

FILING DATE: 13-SEP-1994

APPLICATION NUMBER: US 07/749,611

FILING DATE: 28-AUG-1991

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/576,537

FILING DATE: 31-AUG-1990

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914

REFERENCE/DOCKET NUMBER: 38-21(10660)A

TELECOMMUNICATION INFORMATION:

TELEPHONE: (314)537-6047

TELEFAX: (314)537-6099

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 430 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-476-008-44

QY 1 SGEIDIETIG 11

Db 265 SGEIDVMEARG 275

RESULT 9

US-08-306-063-44

; Sequence 44, Application US/08306063
; Patent No. 5633435

; GENERAL INFORMATION:

; APPLICANT: Barry, Gerard F.

; APPLICANT: Kishore, Ganesh M.

; APPLICANT: Padgett, Stephen R.

; APPLICANT: Stallings, William C.

; TITLE OF INVENTION: Glycosyltransferase-3-Phosphate Synthases

; NUMBER OF SEQUENCES: 69

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. BBAF

STREET: 700 Chesterfield Village Parkway

CITY: St. Louis

STATE: Missouri

COUNTRY: USA

ZIP: 63198

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

Query Match 54.7%; Score 35; DB 1; Length 430;

Best Local Similarity 63.6%; Pred. No. 80;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(10660)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)537-6099
TELEFAX: (314)537-6047
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-306-063-44

Query Match 54.7%; Score 35; DB 1; Length 430;
Best Local Similarity 63.6%; Pred. No. 80;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGEIDIIETIG 11
|||:|:|
DB 265 SGIDIVKMG 275

RESULT 10
US-08-833-485-44
Sequence 44, Application US/08833485
Patent No. 5804425
GENERAL INFORMATION:
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
NUMBER OF SEQUENCES: 5-Enolpyruvylshikimate-3-Phosphate Synthases
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B84F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,485
FILING DATE: 07-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.
REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-485-44

Query Match 54.7%; Score 35; DB 1; Length 430;
Best Local Similarity 63.6%; Pred. No. 80;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGEIDIIETIG 11
|||:|:|
DB 265 SGIDIVKMG 275

RESULT 11
US-09-137-440-44
Sequence 44, Application US/09137440
Patent No. 6248876
GENERAL INFORMATION:
APPLICANT: Barry, Gerard F.
APPLICANT: Kishore, Ganesh M.
APPLICANT: Padgett, Stephen R.
APPLICANT: Stallings, William C.
TITLE OF INVENTION: Glyphosate Tolerant
NUMBER OF SEQUENCES: 5-Enolpyruvylshikimate-3-Phosphate Synthases
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dennis R. Hoerner, Jr., Monsanto Co. B84F
STREET: 700 Chesterfield Village Parkway
CITY: St. Louis
STATE: Missouri
COUNTRY: USA
ZIP: 63198

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/137,440
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/833,485
FILING DATE: 07-APR-1997
APPLICATION NUMBER: US 08/306,063
FILING DATE: 13-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/749,611
FILING DATE: 28-AUG-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/576,537
FILING DATE: 31-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Hoerner Jr., Dennis R.

REGISTRATION NUMBER: 30,914
REFERENCE/DOCKET NUMBER: 38-21(15117)A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (314)737-6099
TELEFAX: (314)737-6047
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-137-440-44

Query Match 54.7%; Score 35; DB 4; Length 430;
Best Local Similarity 63.6%; Pred. No. 80;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SGEIDITE 11
1111111111
265 SGEIDITE 275

RESULT 12
US-08-392-828C-2
Sequence 2, Application US/08392828C
Patent No. 5785962
GENERAL INFORMATION:
APPLICANT: IWANAGA, SADAOKI
APPLICANT: MOTA, TATSUSHI
APPLICANT: SEKI, NORIHI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
STREET: 53 STATE STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/392,828C
FILING DATE: 28-FEB-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, PAULA A
REGISTRATION NUMBER: 32,503
REFERENCE/DOCKET NUMBER: FJN-033
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-392-828C-2

Query Match 53.1%; Score 34; DB 1; Length 654;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDITE 8

Db 116 SGEIDITE 123
1111111111

RESULT 13
US-09-330-945-2
Sequence 2, Application US/09330945
Patent No. 6077946
GENERAL INFORMATION:
APPLICANT: IWANAGA, SADAOKI
APPLICANT: MOTA, TATSUSHI
APPLICANT: SEKI, NORIHI
APPLICANT: ODA, TOSHIO
TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB
TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
STREET: 125 HIGH STREET
CITY: BOSTON
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/330,945
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/119,995
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, EDMUND R
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: FJN-032DV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 654 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-330-945-2

Query Match 53.1%; Score 34; DB 3; Length 654;
Best Local Similarity 87.5%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDITE 8
1111111111
Db 116 SGEIDITE 123

RESULT 14
US-09-396-651B-2
Sequence 2, Application US/09396651B
Patent No. 6225076
GENERAL INFORMATION:
APPLICANT: Darst, Seth A
APPLICANT: Zhang, Gongyi
APPLICANT: Campbell, Elizabeth
APPLICANT: Minakim, Leonid
APPLICANT: Severinov, Konstantin
TITLE OF INVENTION: A CRYSTAL OF BACTERIAL CORE RNA POLYMERASE AND METHODS
TITLE OF INVENTION: OF USE THEREOF

```

: FILE REFERENCE: 600-1-258
: CURRENT APPLICATION NUMBER: US/09/396,651B
: CURRENT FILING DATE: 1999-09-15
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 1119
: TYPE: PRT
: ORGANISM: Thermus aquaticus
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (695)..(696)
: OTHER INFORMATION: Any amino acids can be at these two positions.
US-09-396-651B-2

```

Query Match	53.1%	Score 34:	DB 4:	Length 1119:
Best Local Similarity	54.5%	Pred. No.	3.6e+02:	
Matches	6:	Conservative	3:	Mismatches 2: Indels 0: Gaps 0:

```

      3 EIDIETIGNR 13
      1:11:111
Db    321 EVDDIDHIGNR 331

```

RESULT 15
 5180808-2
 Patent No. 5180808
 APPLICANT: RUOSLAHTI, EERKI I.
 TITLE OF INVENTION: VERISCAN CORE PROTEIN, NUCLEIC ACID
 SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-
 ANTIBODIES, AND METHODS OF DETECTING THE SAME
 NUMBER OF SEQUENCES: 4
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/441,179
 FILING DATE: 27-NOV-1989
 SEQ ID NO.: 2
 LENGTH: 2409
 5180808-2

Query Match	53.1%	Score 34:	DB 6:	Length 2409:
Best Local Similarity	55.6%	Pred. No.	8.7e+02:	
Matches 5, Conservative		Mismatches 4:	Indels 0:	Gaps 0:

QY 1 SGEIDIET 9
|||:::
1446 SGEVDIVDS 1454

Search completed: October 11, 2001, 15:52:04
Job time: 207 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:52:08 ; Search time 15.16 Seconds
(without alignments)
29.375 Million cell updates/sec

Title: US-09-596-101a-1

Sequence: 1 SGEIDITIGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 93435 segs, 34255486 residues

Cal number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40	62.5	782	1	PRTP_HSVT2
2	40	62.5	806	1	MOUSE
3	40	62.5	815	1	MK07_HUMAN
4	40	62.5	1302	1	RPOB_SPTCI
5	40	62.5	1342	1	RPOB_BUCAI
6	40	62.5	1342	1	RPOB_BUCAP
7	40	62.5	1342	1	RPOB_ECOLI
8	40	62.5	1342	1	RPOB_SALTY
9	40	62.5	1343	1	RPOB_HAEIN
10	40	62.5	1392	1	RPOB_NEIMA
11	40	62.5	1392	1	RPOB_NEIMB
12	40	62.5	1655	1	N188_YEAST
13	39	60.9	286	1	GUB_RHOMR
14	39	60.9	540	1	TCPO_CANAL
15	39	60.9	665	1	CNG_DROME
16	38	59.4	371	1	MAIK_ECOLI
17	38	59.4	1103	1	RPOB_CYPAP
18	38	59.4	1143	1	RPOB_PORPU
19	37	57.8	309	1	AGAR_STRCO
20	37	57.8	682	1	EL3B_BACCI
21	37	57.8	1375	1	RPOB_COXBU
22	37	57.8	1492	1	AT7A_RAT
23	36	56.2	446	1	FGIX_COCCA
24	36	56.2	737	1	YMA2_YEAST
25	36	56.2	1096	1	RPOB_GUTTH
26	36	56.2	1194	1	APAE_HUMAN
27	35.5	55.5	546	1	TCPO_SCHPO
28	35	54.7	151	1	VG1A_BP22
29	35	54.7	261	1	YV00_MYCIE
30	35	54.7	313	1	K1P8_HAEIN
31	35	54.7	314	1	MTRH_METMA
32	35	54.7	393	1	DON3_CABEL
33	35	54.7	430	1	ARO4_STRAU

34	35	54.7	546	1	TRCD_ARCFU
35	35	54.7	690	1	CYG3_HUMAN
36	35	54.7	690	1	CYG3_RAT
37	35	54.7	691	1	CYG3_BOVIN
38	35	54.7	728	1	ARPA_ECOLI
39	35	54.7	788	1	SYFB_THERMA
40	35	54.7	897	1	DYHC_HUMAN
41	35	54.7	2511	1	FAS_CHICK
42	35	54.7	4344	1	DYHC_EMENT
43	35	54.7	4349	1	DYHC_FUSO
44	35	54.7	4367	1	DYHC_NEUCR
45	35	54.7	4639	1	DYHC_DROME

ALIGNMENTS

RESULT 1
PRTP_HSVT2
ID PRTP_HSVT2 STANDARD: PRT: 782 AA.
AC O9WR16;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PROBABLE PROCESSING AND TRANSPORT PROTEIN.
GN PRTP.
OS Herpesvirus tupaia (Strain 2) (THV-2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae.
OX NCBI_TaxID=132678;
RN [1]
RX MEDLINE=99319892; PubMed=10392721;
RA Bahr U., Springfield C., Tidona C.A., Darai G.;
RT "Structural organization of a conserved gene cluster of Tupaia
herpesvirus encoding the DNA polymerase, glycoprotein B, a probable
RT processing and transport protein, and the major DNA binding protein.";
RL Virus Res. 60:123-136(1999).
CC -1- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLLOCATION OF THE VIRUS
GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION
(BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF084543; AAD42934.1; -
DR InterPro: IPR000501; -
DR Pfam: PF01366; PRT: 1.
KW Capsid assembly.
SQ SEQUENCE: 782 AA; 87392 MW; 4DE40314E5BABB9B CRC64;

Query Match 62.5%; Score 40; DB 1; Length 782;
Best Local Similarity 65.6%; Pred. No. 17;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 EIDITIGNR 13
DB 236 EDDIVETVGR 246

RESULT 2
MK07_MOUSE
ID MK07_MOUSE STANDARD: PRT: 806 AA.
AC O9WVS8;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)

FT MOD.RES 220 220 PHOSPHORYLATION (ACTIVATES THE KINASE)
 FT CONFLICT 19 32 (BY SIMILARITY)
 FT CONFLICT 609 609 AREGTRPHRCIS -> GPVKEPATASVA (IN
 FT SEQUENCE 815 AA: 88636 MW: 379AD69803207CCF CRC64;
 REF. 2).
 L -> V (IN REF. 2).

Query Match 62.5%; Score 40; DB 1; Length 815;
 Best Local Similarity 80.0%; Pred. No. 18;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 EIDIIETIGN 12
 I:|||||
 DB 53 EYEIIETIGN 62

RESULT 4
 PROB.SPICI STANDARD; PRT; 1302 AA.

DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).

GN PROB.
 OS Spiroplasma citri.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;
 OC Spiroplasmataceae; Spiroplasma.
 OX NCBI_TaxID=2133;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-ROZHP;
 RX MEDLINE=96257200; PubMed=8675039;
 LAIget F., Gaurivaud P., Bove J.;
 RT "The unique organization of the rpoB region of Spiroplasma citri: a
 restriction and modification system gene is adjacent to rpoB";
 RL Gene 171:95-98(1996).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
 RNA(N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 BETA CHAIN.

-1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: U25815; AAC44217.1; -
 CC InterPro: IPR001572; -
 DR Pfam: PF00562; RNA_POL_B; 1.
 DR PROSITE: PS01166; RNA_POL_BETA; 1.
 KM Transferrase; Transcription; DNA-directed RNA polymerase.
 SO SEQUENCE 1302 AA; 146533 MW; 8D971C23EB9FEB2F CRC64;

Query Match 62.5%; Score 40; DB 1; Length 1302;
 Best Local Similarity 58.3%; Pred. No. 30;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGN 13
 I:|:|:|
 DB 475 GEVDIDHIGNR 486

RESULT 5
 PROB.BUCAP STANDARD; PRT; 1342 AA.

DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).

GN PROB OR B0034.
 OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum
 symbiotic bacterium).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=118099;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-TOKYO 1998;
 RX MEDLINE=20445173; PubMed=10993077;
 RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
 RT "Genome sequence of the endocellular bacterial symbiont of aphids
 Buchnera sp. Aps.";
 RL Nature 407:81-86(2000).

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +
 RNA(N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 BETA CHAIN.

-1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: AP001118; BAB12761.1; -
 DR PROSITE: PS01166; RNA_POL_BETA; 1.
 KM Transferrase; Transcription; DNA-directed RNA polymerase.
 SO SEQUENCE 1342 AA; 151467 MW; 95252459873DF940 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIIETIGNR 13
 I:|:|:|
 DB 440 GEVDIDHIGNR 451

RESULT 6
 PROB.BUCAP STANDARD; PRT; 1342 AA.

DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 BETA CHAIN) (RNA POLYMERASE BETA SUBUNIT).

GN PROB.
 OS Buchnera aphidicola (subsp. Schizaphis graminum).
 OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
 OX NCBI_TaxID=98794;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=93160925; PubMed=1369199;
 RA Clark M.A., Baumann L., Baumann P.;

CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
 CC -1- CAUTION: REF.6 SEQUENCE DIFFERS FROM THAT SHOWN EXTENSIVELY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: V00339; CAA23625.1; -
 DR EMBL: U00006; AAC43085.1; -
 DR EMBL: A0000472; AAC76961.1; -
 DR EMBL: U76222; AAB18647.1; -
 DR EMBL: V00340; CAA23627.1; -
 DR EMBL: V00341; CAA23629.1; -
 DR EMBL: M38303; AAA24583.1; -
 DR EMBL: M38304; AAA24580.1; -
 DR EMBL: M38292; AAA24579.1; -
 DR EMBL: M38293; AAA24581.1; -
 DR EMBL: M38287; AAA24585.1; -
 DR EMBL: U77436; AAD09605.1; -
 DR PIR: A00688; RNECB.
 DR PIR: A00689; A00689.
 DR SWISS-2DPAGE: P00575; COLI.
 DR ECODBASE: D157.0; 6TH EDITION.
 DR Ecogene: EG10894; rpoB.
 DR InterPro: IPR001572; -
 DR Pfam: PF00562; RNA_POL_B; 1.
 DR PROSITE: PS01166; RNA_POL_BETA; 1.
 KM Transference; transcription; DNA-directed RNA polymerase.
 FT MTAGEN 813 813 E->K: DISRUPT THE ENZYME'S ACTIVE CENTER.
 FT CONFLICT 106 107 ER->G (IN REF. 5).
 FT CONFLICT 384 391 LEFNLFFS->CSRTCSPT (IN REF. 5).
 FT CONFLICT 516 516 D->V (IN REF. 1 AND 9).
 SQ SEQUENCE 1342 AA; 150631 MW; F9E95344C54AB18 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIGNR 13
 ||:|:|:|
 440 GEYDDIDHLGNR 451

RESULT 8
 RPOB_SALTY STANDARD; PRT; 1342 AA.
 AC P06173;
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE BETA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 DE DNA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
 GN RPOB.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Salmonella.
 OX NCBI_Taxid=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sverdlov E.D., Listitsyn N.A., Guryev S.O., Monastyrskaya G.S.;
 RT "Nucleotide sequence of the rpoB gene of Salmonella typhimurium coding
 RT for the beta subunit of RNA polymerase.";
 RL Dokl. Biochem. 287:62-65(1986).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89052707; PubMed=3056723;

RA Listitsyn N.A., Monastyrskaya G.S., Sverdlov E.D.;
 RT "Genes coding for RNA polymerase beta subunit in bacteria.
 RT structure/function analysis.";
 RL Eur. J. Biochem. 177:363-369(1988).
 RN [3]
 RP SEQUENCE OF 187-1144 FROM N.A.
 RX MEDLINE=86269165; PubMed=3015156;
 RA Sverdlov E.D., Listitsyn N.A., Guryev S.O., Smitnov V.V.,
 RA Gostapshov V.M., Monastyrskaya G.S.;
 RT "Genes encoding the beta subunit of bacterial RNA-polymerases. I.
 RT Primary structure of the EcoRI-C fragment of the Salmonella
 RT typhimurium gene rpoB.";
 RL Bioorg. Khim. 12:693-707(1986).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE - N PYROPHOSPHATE +
 CC RNA(N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X04642; CAA28302.1; -
 DR EMBL: M38311; AAA27215.1; -
 DR PIR: S01794; RNEBET.
 DR StyGene: SG10368; rpoB.
 DR InterPro: IPR001572; -
 DR Pfam: PF00562; RNA_POL_B; 1.
 DR PROSITE: PS01166; RNA_POL_BETA; 1.
 KM Transference; transcription; DNA-directed RNA polymerase.
 FT CONFLICT 401 401 G->A (IN REF. 3).
 SQ SEQUENCE 1342 AA; 150557 MW; EDBF979D3E9E4D8 CRC64;

Query Match 62.5%; Score 40; DB 1; Length 1342;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIGNR 13
 ||:|:|:|
 Db 440 GEYDDIDHLGNR 451

RESULT 9
 RPOB_HAEIN STANDARD; PRT; 1343 AA.
 AC P43738;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE BETA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6) (TRANSCRIPTASE
 DE DNA CHAIN) (RNA POLYMERASE BETA SUBUNIT).
 GN RPOB OR H10515.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OX NCBI_Taxid=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,


```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; 254353; CAA91164.1; -
DR EMBL; AE002371; AAF40591.1; ALT_INIT.
DR TIGR; NMB0132; -
DR InterPro: IPR001572; -
DR Pfam: PF00562; RNA_POL_B; 1.
DR PROSITE; PS01166; RNA_POL_BETA; 1.
KW Translasease; Transcription; DNA-directed RNA polymerase.
FT CONFLICT 2 28 N -> S (IN REF. 1).
FT CONFLICT 28 28 L -> I (IN REF. 1).
FT CONFLICT 224 252 ILIDFDEKFEYLSNGYQDTLVADRLKG -> NLGYFLRQ
RVNLFVPRKCSNREGRRSES (IN REF. 1).
FT CONFLICT 284 284 N -> L (IN REF. 1).
FT CONFLICT 350 360 AYISNTLTDE -> VISPICVRMK (IN REF. 1).
FT CONFLICT 378 378 MISSING (IN REF. 1).
FT CONFLICT 648 648 A -> G (IN REF. 1).
FT CONFLICT 706 706 A -> P (IN REF. 1).
FT CONFLICT 717 718 VP -> SA (IN REF. 1).
FT CONFLICT 748 749 GG -> A (IN REF. 1).
FT CONFLICT 818 818 F -> L (IN REF. 1).
FT CONFLICT 836 837 GY -> VN (IN REF. 1).
FT CONFLICT 1190 1192 YNG -> SR (IN REF. 1).
FT CONFLICT 1201 1201 A -> S (IN REF. 1).
FT CONFLICT 1244 1246 DDP -> EDA (IN REF. 1).
FT CONFLICT 1373 1373 F -> L (IN REF. 1).
SQ SEQUENCE 1392 AA; 155708 MW; 92AA331A597898F1 CRC64;

Query Match
Best Local Similarity 62.5%; Score 40; DB 1; Length 1392;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 GEIDIETIGNR 13
DQ 467 GEVDIDHIGNR 478
DQ 1188 YEAST STANDARD; PRT; 1655 AA.
P52593;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE NCICLPOFORIN NUP188 (NUCLEAR PORE PROTEIN NUP188).
GN NUP188 OR YML103C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RX Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RA Zabel U., Doye V., Tekotte H., Wepf R., Hurt E.C.;
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN 12
RP SEQUENCE FROM N.A.
RA Neirbass U., Rott M.P., Maguire S., Blobel G., Wozniak R.W.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN 13
RP SEQUENCE FROM N.A.
RA STAIN-S288C / AB972;
RA Devlin K., Churcher C., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN 14
RP SEQUENCE OF 1187-1205 AND 1611-1629.

```

```

RX MEDLINE=96095775; PubMed=8525278;
RA Aitchison J.D., Rott M.P., Marelli M., Blobel G., Wozniak R.W.;
RT "Two novel related yeast nucleoporins Nup170p and Nup157p:
RT complementation with the vertebrate homologue Nup155p and functional
RT interactions with the yeast nuclear pore-membrane protein Pom152p.";
RL J. Cell Biol. 131:1133-1148(1995).
CC - FUNCTION: INTERACTS WITH THE CORE STRUCTURE OF THE NUCLEAR PORE
CC COMPLEX (NPC). MAY PROVIDE THE NECESSARY ASYMMETRY REQUIRED FOR
CC ANCHORING STRUCTURES SUCH AS CYTOSOLIC FILAMENTS AND THE
CC NUCLEOLAR PLASMA CAGE.
CC - SUBUNIT: INTERACTS WITH NUP170 AND/OR POM152.
CC - SUBCELLULAR LOCATION: NUCLEAR PORE COMPLEX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X90580; CAA62208.1; -
DR EMBL; U47107; AAA88904.1; -
DR EMBL; X80835; CAA56794.1; -
DR SGD; S0004571; NUP188.
KW Nuclear protein; Transport.
SQ SEQUENCE 1655 AA; 188576 MW; B590051EE60464C CRC64;

Query Match
Best Local Similarity 72.7%; Score 40; DB 1; Length 1655;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 GSEIDIETIG 11
DQ 1290 GSEIDFIRKNG 1300
DQ 1188 YEAST STANDARD; PRT; 286 AA.
P45798;
AC GUB_RHOMR
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE BETA-GLUCANASE PRECURSOR (EC 3.2.1.73) (ENDO-BETA-1,3-1,4 GLUCANASE)
DE (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE) (LICHENASE).
GN BGIA.
OS Rhodothermus marinus.
OC Bacteria; CFM group; Rhodothermus group; Rhodothermus.
OX NCBI_TaxID=29549;
RX MEDLINE=95010084; PubMed=7925416;
RA Spillier R., Hregyrdsson G.O., Kristjansson J.K.;
RA Eggertsson G., Palsdottir A.;
RT Cloning and sequencing of a Rhodothermus marinus gene, bgia, coding
RT for a thermostable beta-glucanase and its expression in Escherichia
RT coli.";
RL Eur. J. Biochem. 224:923-930(1994).
CC - FUNCTION: SHOWS ACTIVITY ON LICHENAN, BETA-GLUCAN AND LAMINARIN
CC BUT NOT ON CMC CELLULOSE OR XYLAN. THE ENZYME HAS A TEMPERATURE
CC OPTIMUM OF 85 DEGREES CELSIUS AND A PH OPTIMUM OF 7.0.
CC - CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLYCOSIDIC
CC LINKAGES IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.
CC - SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: 004836; AAG0459.1; -
 CC HSSP: P23904; IADR.
 DR InterPro: IPR000757; -
 DR Pfam: PF00722; Glyco_hydro.16; 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 KM Hydrolyase; Glycosidase; Signal.
 FT SIGNAL 1 30
 FT CHAIN 31 286
 FT ACT_SITE 158 158
 FT ACT_SITE 163 163
 FT SEQUENCE 286 AA; 33145 MW; 7215C33624135191 CRC64;

Query Match 60.9%; Score 39; DB 1; Length 286;
 Best Local Similarity 63.6%; Pred. No. 9.1;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 SEIDITITIG 11
 :|||||:1
 156 NGSDIMEHYG 166

RESULT 14
 TCPO_CANAL STANDARD; PRT; 540 AA.
 AC P47828;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DE T-COMPLEX PROTEIN 1, THERA SUBUNIT (TcP-1-THERA) (CCT-THERA).
 GN CCT8.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 ON NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-ATCC 10231;
 RX MEDLINE=96367593; PubMed=8771707;
 RA Stoldt V., Rademacher F., Kehren V., Ernst J.F., Sherman F.;
 RT "Review: the Cct eukaryotic chaperonin subunits of Saccharomyces
 cerevisiae and other yeasts.";
 RL Yeast 12:523-529(1996).
 RN [2]
 RP REVISIONS.
 RP STRAIN-ATCC 10231;
 Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 Ernst J.F.;
 -1- FUNCTION: MOLECULAR CHAPERONE; ASSIST THE FOLDING OF PROTEINS UPON
 ATP HYDROLYSIS. KNOWN TO PLAY A ROLE, IN VITRO, IN THE FOLDING OF
 ACTIN AND TUBULIN. IN YEAST MAY PLAY A ROLE IN MITOTIC SPINDLE
 FORMATION (BY SIMILARITY).
 -1- SUBUNIT: HETERO-OLIGOMERIC COMPLEX OF ABOUT 850 TO 900 KDA THAT
 FORMS TWO STACKED RINGS, 12 TO 16 NM IN DIAMETER (BY SIMILARITY).
 -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 -1- SIMILARITY: BELONGS TO THE TcP-1 CHAPERONIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR PROSITE: PS00751; TCPL2; 1.
 DR PROSITE: PS00995; TCPL3; FALSE_NEG.
 KM Chaperone ATP-binding; Multigene family.
 SQ SEQUENCE 540 AA; 58918 MW; F962285DA6EB03DE CRC64;

Query Match 60.9%; Score 39; DB 1; Length 540;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 GEIDITET 9
 :|||||
 Db 342 GEIDITET 349

RESULT 15
 CNG_DROME STANDARD; PRT; 665 AA.
 ID CNG_DROME
 AC 024278;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE CYCLIC-NUCLEOTIDE-GATED CATION CHANNEL (CNG CHANNEL).
 GN CNG.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 ON NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95045396; PubMed=7957070;
 RA Baumann A., Frings S., Godde M., Selfert R., Kaupp U.B.;
 RT "Primary structure and functional expression of a Drosophila cyclic
 nucleotide-gated channel present in eyes and antennae.";
 RL EMBO J. 13:5040-5050(1994).
 CC -1- FUNCTION: APPROXIMATELY 50-FOLD MORE SENSITIVE TO CGMP THAN TO
 CAMP. MAY BE INVOLVED IN TRANSDUCTION CASCADES OF BOTH
 INTRACELLULAR PHOTORECEPTORS AND OLFACTORY SENSILLAE.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTENNAE AND THE VISUAL SYSTEM.
 CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL
 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

FT	DOMAIN	503	665	CYTOPLASMIC (POTENTIAL).
FT	NP BIND	437	559	CAMP (BY SIMILARITY).
FT	BINDING	496	496	CAMP (POTENTIAL).
FT	BINDING	511	511	CAMP (POTENTIAL).
FT	CARBOHYD	135	135	N-LINKED (GLCNAC. . .) (POTENTIAL).
SO	SEQUENCE	665 AA;	75922 MM;	9F1BDC5D9581C8DB CRC64;

Query Match 60.9%; Score 39; DB 1; Length 665;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 2 GEIDIERIGNR 13
 ||: :|| |||
 Db 495 GEVSYLEIAGNR 506

Search completed: October 11, 2001, 15:57:05
 Time: 297 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:51:42 ; Search time 34.86 Seconds
(Without alignments)
49.339 Million cell updates/sec

Title: US-09-596-101A-1
Perfect score: 64
Sequence: 1 SGEIDIIETIGNR 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:REMBL.16:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organelle:*
10: sp.phage:*
11: sp.plant:*
12: sp.rodent:*
13: sp.unclassified:*
14: sp.vertebrate:*
15: sp.virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	100.0	384	5	077072 elsenia foe
2	50	78.1	306	2	051333
3	47	73.4	361	5	09U064
4	45	70.3	294	2	007242
5	45	70.3	499	5	026660
6	44	68.8	422	2	09EWR5
7	44	68.8	435	2	068641
8	43	67.2	285	2	09E3A0
9	43	67.2	1383	2	09KJMS
10	43	67.2	1383	2	09KJMS
11	41	64.1	736	10	09SVX7
12	40	62.5	342	2	09FA06
13	40	62.5	342	2	09FA05
14	40	62.5	342	2	09FA04
15	40	62.5	342	2	09E075
16	40	62.5	342	2	09E076
17	40	62.5	342	2	09E076
18	40	62.5	395	5	017492
19	40	62.5	806	11	09WVF4

20	40	62.5	1342	2	09L9J8	091938 salmonella
21	40	62.5	1375	2	09K930	09K930 vibrio chol
22	39	60.9	276	2	052734	052734 rhodotherm
23	39	60.9	642	2	09WXX1	09WXX1 thermotoga
24	39	60.9	646	2	060039	060039 thermotoga
25	39	60.9	665	5	09V7L5	09V7L5 drosophila
26	39	60.9	967	10	09LXV7	09LXV7 arabidopsis
27	39	60.9	1324	2	059328	059328 clostridium
28	38	59.4	263	2	09WZC8	09WZC8 thermotoga
29	38	59.4	297	1	073951	073951 pyrococcus
30	38	59.4	298	2	P73049	P73049 synecocyst
31	38	59.4	544	2	050223	050223 thiochalcit
32	38	59.4	608	14	09YVY2	09YVY2 melanoplus
33	37	57.8	316	10	09WZB6	09WZB6 arabidopsis
34	37	57.8	664	2	09K920	09K920 bacillus ha
35	37	57.8	794	2	09RIA3	09RIA3 streptomyce
36	37	57.8	877	2	045095	045095 bacillus ci
37	37	57.8	908	1	028789	028789 archaeoglob
38	37	57.8	1292	5	09Y0C8	09Y0C8 drosophila
39	37	57.8	1404	5	045251	045251 caenorhabdi
40	37	57.8	1429	5	015764	015764 tetrahymena
41	36.5	57.0	188	1	09Y9P2	09Y9P2 aeropyrum p
42	36	56.2	88	1	09HS27	09HS27 halobacteri
43	36	56.2	223	1	027105	027105 methanobact
44	36	56.2	310	2	09KNJ8	09KNJ8 vibrio chol
45	36	56.2	352	5	09V4R3	09V4R3 drosophila

ALIGNMENTS

RESULT 1
077072 PRELIMINARY; PRT; 384 AA.
ID 077072;
AC 077072;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE COELOMIC CYTOLYTIC FACTOR 1.
GN CCF1.
OS Elsenia foetida (Common branding worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Elsenia.
OX NCBI_TaxID=6396;
RX MEDLINE=98406152; PubMed=9733802;
RA Beschta A., Bilej M., Hanssens F., Raymakers J., Van Dyck E.,
RA Reyets H., Blys L., Gomez J., De Baetselier P., Timmermans M.,
RT Identification and cloning of a glucan- and lipopolysaccharide-
RT binding protein from Elsenia foetida earthworm involved in the
RT activation of prophenoloxidase cascade.*;
RL J. Biol. Chem. 273:24948-24954(1998).
DR EMBL: AF030028; AAC35887.1; -
SQ SEQUENCE 384 AA; 44322 MW; C90B5C94003BAD6D CRC64;

Query Match 100.0%; Score 64; DB 5; Length 384;

Best Local Similarity 100.0%; Pred. No. 0.0028;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SGEIDIIETIGNR 13
Db 178 SGEIDIIETIGNR 190

RESULT 2
ID 051333 PRELIMINARY; PRT; 306 AA.
AC 051333;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE BETA-1,3-GLUCANASE IIA.
 GN BGLIIA.
 OS Oerskovia xanthineolytica.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Micrococcales; Cellulomonadaceae; Cellulomonas.
 CC NCBI_TaxID=1826;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LL 6109.
 RX MEDLINE=96345651; PubMed=8755914;
 RA Ferrer P., Halkier T., Hedegaard L., Savva D., Diers I., Asenjo J.A.;
 RT "Nucleotide sequence of a beta-1,3-glucanase isoenzyme IIA gene of
 RT Oerskovia xanthineolytica LL 6109 (Cellulomonas cellulans) and initial
 RT characterization of the recombinant enzyme expressed in Bacillus".
 RT subills."
 RL J. Bacteriol. 178:4751-4757(1996).
 DR EMBL: U56935; AAC4371.1; -.
 DR HSSP: P23904; IAJK.
 DR InterPro: IPR000757; -.
 DR Pfam: PF00722; Glyco_hydro.16; 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 CA SEQUENCE 306 AA; 32835 MW; CDBDB8C1A6F8DC04 CRC64;

Query Match 78.1%; Score 50; DB 2; Length 306;
 Best Local Similarity 75.0%; Pred. No. 0.69;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGEIDIIETIGN 12
 DB 180 SGEIDIMENVGN 191

RESULT 3
 ID 090064 PRELIMINARY; PRT; 361 AA.
 AC 090064;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE LIPOPOLYSACCHARIDE AND BETA-1,3-GLUCAN BINDING PROTEIN PRECURSOR.
 GN LGBP.
 OS Pacifastacus leniusculus (Signal crayfish).
 CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 CC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
 CC Astacoidea; Astacidae; Pacifastacus.
 CC NCBI_TaxID=6720;
 RX [1]
 RP SEQUENCE FROM N.A.
 RS TISSUE=HEMOCYTES;
 MEDLINE=20092910; PubMed=10625682;
 RA Lee S., Wang R., Soderhall K.;
 RT "A lipopolysaccharide- and beta-1,3-glucan-binding protein from
 RT hemocytes of the freshwater crayfish Pacifastacus leniusculus:
 RT purification, characterization, and cDNA cloning.";
 RL J. Biol. Chem. 275:1337-1343(2000).
 DR EMBL: AJ250128; CAB65353.1; -.
 KM Signal.
 FT SIGNAL
 FT CHAIN 16 15 LIPOPOLYSACCHARIDE AND BETA-1,3-GLUCAN
 FT BINDING PROTEIN.
 SO SEQUENCE 361 AA; 41043 MW; 93A5E67911ED6619 CRC64;

Query Match 73.4%; Score 47; DB 5; Length 361;
 Best Local Similarity 75.0%; Pred. No. 2.9;
 Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGEIDIIETIGN 12
 DB 172 SGEIDIVSRGN 183

RESULT 4
 ID 007242 PRELIMINARY; PRT; 294 AA.
 AC 007242;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)
 DE PROBABLE BETA-1,3-GLUCANASE.
 GN RV0315 OR MYC63.20.
 OS Mycobacterium tuberculosis.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=9625987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 DR EMBL: Z96800; CAB09586.1; -.
 DR TuberculList: RV0315; -.
 DR InterPro: IPR000757; -.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; UNKNOWN.1.
 SQ SEQUENCE 294 AA; 32186 MW; 7B9897BA6740BE80 CRC64;

Query Match 70.3%; Score 45; DB 2; Length 294;
 Best Local Similarity 75.0%; Pred. No. 5.1;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SGEIDIIETIGN 12
 DB 169 SGEIDILEWIGN 180

RESULT 5
 ID 026660 PRELIMINARY; PRT; 499 AA.
 AC 026660;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE BETA 1,3-GLUCANASE.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 CC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 CC Echinoida; Echinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 CC Strongylocentrotus.
 CC NCBI_TaxID=7668;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96270625; PubMed=8692900;
 RA Bachman E.S., McClay D.R.;
 RT "Molecular cloning of the first metazoan beta-1,3 glucanase from eggs
 RT of the sea urchin Strongylocentrotus purpuratus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:6808-6813(1996).
 DR EMBL: U49711; AAC47233.1; -.
 SO SEQUENCE 499 AA; 55274 MW; D863F36E0CF55AC CRC64;

Query Match 70.3%; Score 45; DB 5; Length 499;
 Best Local Similarity 66.7%; Pred. No. 9.5;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SGEIDIIETIGN 12

DB 326 SGEIDLVESRGN 337

RESULT 6
ID 09EMR5 PRELIMINARY; PRT; 422 AA.

AC 09EMR5; 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
RT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE PUTATIVE SECRETED HYDROLASE.
GN 3SC660.19.

OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;

RA SEQUENCE FROM N.A.
[1]
STRAIN-A3(2);
Seeger K.J., Harris D.;
Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RA SEQUENCE FROM N.A.
[2]
STRAIN-A3(2);
Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

RA SEQUENCE FROM N.A.
[3]
STRAIN-A3(2);
MEDLINE=9700351; PubMed=8843436;
Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
Kinashi H., Hopwood D.A.;
"A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
Mol. Microbiol. 21:77-96(1996).
EMBL: AL45945; CAC14352.1; -.

RA Hydrolase. 422 AA; 44143 MW; 7105047BE26F3020 CRC64;
SQ SEQUENCE

Query Match 68.8%; Score 44; DB 2; Length 422;
Best Local Similarity 72.7%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIETIG 11
170 SGEIDMENVG 180

RESULT 7
ID 068641 PRELIMINARY; PRT; 435 AA.
AC 068641; 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, last sequence update)
RT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE BETA-1,3-GLUCANASE II.
GN BGLII.
OS Oerskovia xanthineolytica.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococccinae; Cellulomonadaceae; Cellulomonas.
OX NCBI_TaxID=1826;
RA SEQUENCE FROM N.A.
[1]
STRAIN-LL G109;
Ventom A.M., Asenjo J.A.;
Enzyme Microb. Technol. 13:71-75(1991).
[2]
SEQUENCE FROM N.A.
STRAIN-LL G109;
MEDLINE=96409238; PubMed=8814220;
Parrado J., Escuredo P.R., Conejero-Lara F., Kotik M., Ponting C.P.,

RA Asenjo J.A., Dobson C.M.;
RT "Molecular characterisation of a thermostable beta-1,3-glucanase from
Oerskovia xanthineolytica."
RT Blochim. Biophys. Acta 1296:145-151(1996).
[3]

RA SEQUENCE FROM N.A.
[3]
STRAIN-LL G109;
MEDLINE=96250169; PubMed=8659924;
RA Ferrer P., Hedegaard L., Halkier T., Diers I., Savva D., Asenjo J.A.;
RT "Molecular cloning of a lytic beta-1,3-glucanase gene from Oerskovia
xanthineolytica lig109. A beta-1,3-glucanase able to selectively
penetrate the yeast cell wall."
RT Ann. N. Y. Acad. Sci. 782:555-566(1996).
[4]

RA SEQUENCE FROM N.A.
[4]
STRAIN-LL G109;
Ferrer P., Andrews B.A., Asenjo J.A., Hedegaard L., Diers I.;
Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

RA EMBL: AF052745; AAC38290.1; -.
DR HSP; P23904; IAKK.
DR InterPro: IPR000757; -.
DR InterPro: IPR000772; -.

DR Pfam: PF00652; Rich. B. lectin; 1.
DR Pfam: PF00722; Glyco_hydro_16; 1.
DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
DR SMART: SM00458; RICIN; 1.
SQ SEQUENCE 435 AA; 46097 MW; 00F087BE64C0F58 CRC64;

Query Match 68.8%; Score 44; DB 2; Length 435;
Best Local Similarity 72.7%; Pred. No. 12;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 SGEIDIETIG 11
169 SGEIDMENVG 179

RESULT 8
ID 09F3A0 PRELIMINARY; PRT; 285 AA.
AC 09F3A0; 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
RT 01-MAR-2001 (TREMBlrel. 16, last annotation update)
DE PUTATIVE SECRETED HYDROLASE.
GN SC5F1.29C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RA SEQUENCE FROM N.A.
[1]
STRAIN-A3(2);
Saunders D.C., Harris D.;
Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
STRAIN-A3(2);
Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
STRAIN-A3(2);
MEDLINE=9700351; PubMed=8843436;
Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
Kinashi H., Hopwood D.A.;
"A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
Mol. Microbiol. 21:77-96(1996).
EMBL: AL450165; CAC16455.1; -.
KW Hydrolase. 285 AA; 30479 MW; 6977A032A1A31451 CRC64;
SQ SEQUENCE

Query Match 67.2%; Score 43; DB 2; Length 285;

Best Local Similarity 63.6%; Pred. No. 11; Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 SGEIDIIETIGR 11
|||||:|:|:|
DB 160 SGEIDVMEYNG 170

RESULT 9

Q9KJMS PRELIMINARY; PRT; 1383 AA.

AC 09KJMS: 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
RNA POLYMERASE BETA SUBUNIT.

GN RPOB.
Bartonella quintana (Rochalimaea quintana).
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Bartonellaceae; Bartonella.

NCBI_TaxID=803;

SEQUENCE FROM N.A.

RC STRAIN-FULLER.
Renesto P., Gouvernet J., Drancourt M., Raoult D.;

"New approach to identifying Bartonella species on the basis of the
rpoB gene sequencing."

Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
SUBSTRATES (BY SIMILARITY).

-1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE + N PYROPHOSPHATE +
RNA(N) (BY SIMILARITY).

-1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

EMBL; AF165994; AAF80850.1; -

InterPro: IPR001572; -

DR Pfam: PF00562; RNA_POL_B; 1.

DR PROSITE: PS01166; RNA_POL_BETA; 1.

KW DNA-directed RNA polymerase; Transcription; Transferase.
SQ SEQUENCE 1383 AA; 154780 MW; 9B007A67C872498E CRC64;

2 GEIDIIETIGR 13
|||||:|:|:|

DB 454 GEIDIDIDNIGR 465

RESULT 10

Q9KJG4 PRELIMINARY; PRT; 1383 AA.

AC 09KJG4: 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
RNA POLYMERASE BETA SUBUNIT.

GN RPOB.
Bartonella henselae (Rochalimaea henselae).

Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Bartonellaceae; Bartonella.

NCBI_TaxID=38323;

SEQUENCE FROM N.A.

RC STRAIN-HOUSTON-1.
Renesto P., Gouvernet J., Drancourt M., Raoult D.;

"New approach to identifying Bartonella species on the basis of the
rpoB gene sequencing."

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
SUBSTRATES (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE + N PYROPHOSPHATE +
RNA(N) (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

EMBL; AF171070; AAF87049.1; -

DR InterPro: IPR001572; -

DR Pfam: PF00562; RNA_POL_B; 1.

DR PROSITE: PS01166; RNA_POL_BETA; 1.

KW DNA-directed RNA polymerase; Transcription; Transferase.
SQ SEQUENCE 1383 AA; 154852 MW; 2CE87E06A3207B6 CRC64;

QY 2 GEIDIIETIGR 13
|||||:|:|:|

DB 454 GEIDIDIDNIGR 465

RESULT 11

Q9SVX7 PRELIMINARY; PRT; 736 AA.

AC 09SVX7: 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)

DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE HYPOTHEORETICAL 84.2 KDA PROTEIN.

GN f15B.180.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;

OC Brassicales; Brassicaceae; Arabidopsis.

NCBI_TaxID=3702.

SEQUENCE FROM N.A.

RA Benes V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W.,

RA Mayer K.F.X., Lemcke K., Scheller C., Queller F., Salanoubat M.;

RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.

RA EU Arabidopsis sequencing project;

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL049660; CAB41192.1; -

DR InterPro: IPR000561; -

DR PRINTS: PR00011; EGFAMININ.

DR PROSITE: PS00022; EGF_1; UNKNOWN_2.

DR PROSITE: PS01186; EGF_2; 2.

KW EGF-like domain; Glycoprotein; Hypothetical protein.

SQ SEQUENCE 736 AA; 84202 MW; 349E0F1E6A28C9A CRC64;

QY 3 EIDIIETIGR 13
|||||:|:|:|

DB 109 EVDIIETIGR 119

RESULT 12

Q9F406 PRELIMINARY; PRT; 342 AA.

AC 09F406: 01-MAR-2001 (TReMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)

DE 01-MAR-2001 (TReMBLrel. 16, Last annotation update)

DE RNA POLYMERASE BETA SUBUNIT (FRAGMENT).
 GN RPOB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B:4:PL.10;
 RA Stefanelli P., Fazio C., La Rosa G., Marianelli C., Muscillo M.,
 RT "Rifampin-resistant meningococci causing invasive disease: detection
 of point mutations in the rpoB gene and molecular characterization of
 the strain."
 RL J. Antimicrob. Chemother. 0:0-0(0).
 DR EMBL: AJ270494; CAC14798.1; -.
 FT NON_TER 1 342
 FT SEQUENCE 342 AA; 37237 MW; E4C6B3719348B0AB CRC64;
 SQ

Query Match 62.5%; Score 40; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIETIGNR 13
 ||| | : |||
 DB 36 GEVDDIDHGNR 47

RESULT 13

O9F405 PRELIMINARY; PRT; 342 AA.
 AC O9F405;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE RNA POLYMERASE BETA SUBUNIT (FRAGMENT).
 GN RPOB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B:4:PL.10;
 RA Stefanelli P., Fazio C., La Rosa G., Marianelli C., Muscillo M.,
 RT "Rifampin-resistant meningococci causing invasive disease: detection
 of point mutations in the rpoB gene and molecular characterization of
 the strain."
 RL J. Antimicrob. Chemother. 0:0-0(0).
 DR EMBL: AJ270501; CAC14805.1; -.
 FT NON_TER 1 342
 FT SEQUENCE 342 AA; 37268 MW; 5C08CDA6C25AFBCF CRC64;
 SQ

Query Match 62.5%; Score 40; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIETIGNR 13
 ||| | : |||
 DB 36 GEVDDIDHGNR 47

RESULT 14

O9F404 PRELIMINARY; PRT; 342 AA.
 AC O9F404;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE RNA POLYMERASE BETA SUBUNIT (FRAGMENT).

GN RPOB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B:4:PL.10;
 RA Stefanelli P., Fazio C., La Rosa G., Marianelli C., Muscillo M.,
 RT "Rifampin-resistant meningococci causing invasive disease: detection
 of point mutations in the rpoB gene and molecular characterization of
 the strain."
 RL J. Antimicrob. Chemother. 0:0-0(0).
 DR EMBL: AJ270502; CAC14806.1; -.
 FT NON_TER 1 342
 FT SEQUENCE 342 AA; 37264 MW; D04BAA7653E5B324 CRC64;
 SQ

Query Match 62.5%; Score 40; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIETIGNR 13
 ||| | : |||
 DB 36 GEVDDIDHGNR 47

RESULT 15

O9E075 PRELIMINARY; PRT; 342 AA.
 AC O9E075;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DE 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE RNA POLYMERASE BETA SUBUNIT (FRAGMENT).
 GN RPOB.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-B:14:PL.12, C:2B:PL.2, C:2B, AND:PL.10;
 RA Stefanelli P., Fazio C., La Rosa G., Marianelli C., Muscillo M.,
 RT "Rifampin-resistant meningococci causing invasive disease: detection
 of point mutations in the rpoB gene and molecular characterization of
 the strain."
 RL J. Antimicrob. Chemother. 0:0-0(0).
 DR EMBL: AJ270504; CAC14808.1; -.
 DR EMBL: AJ270500; CAC14804.1; -.
 DR EMBL: AJ270503; CAC14807.1; -.
 FT NON_TER 1 342
 FT SEQUENCE 342 AA; 37253 MW; 2CD9BD9709A5DEDE CRC64;
 SQ

Query Match 62.5%; Score 40; DB 2; Length 342;
 Best Local Similarity 58.3%; Pred. No. 48;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 2 GEIDIETIGNR 13
 ||| | : |||
 DB 36 GEVDDIDHGNR 47

Search completed: October 11, 2001, 15:56:43
 Job time: 301 sec

DR MPI: 1999-385905/32.
DR N-PSDB: AAX83611.

XX Eissenia foetida polypeptides derived from coelomic cytolytic factor 1
XX Claim 2; Page 48-49; 49pp; English.

CC The present sequence represents the Eissenia foetida coelomic cytolytic
CC factor 1 (CCF-1). The protein has antiparasitic, antibacterial and
CC antiinflammatory activity. Recombinant coelomic cytolytic factor 1
CC (rCCF-1) is trypanolytic for the African trypanosome Trypanosoma brucei
CC in a dose-dependent manner. The trypanolytic activity of rCCF-1 can be
CC inhibited by anti-CCF-1 and anti-tumour necrosis factor (TNF)/TNF-
CC monoclonal antibodies. Furthermore, N,N'-diacetylethanolamine inhibits
CC potentially trypanolytic activity of rCCF-1. These data corroborate the
CC findings that CCF-1 shares a trypanolytic, lectin-like domain with
CC TNF-alpha. CCF-1 is useful to treat trypanosomal or bacterial infections
CC or cancer. The proteins and peptides are also useful in tumour therapy,
CC inflammation and other areas of immunology. The amino acid sequence of
CC derived from a 4.42 kDa cytolytic protein named CCF-1 that binds
CC lipopolysaccharide and beta-1,3-glucan. The factor resembles the
CC vertebrate tumour necrosis factor-alpha (TNF-alpha), and may be used as
CC an alternative for TNF-alpha.

XX Sequence 384 AA;

Query Match 100.0%; Score 2240; DB 20; Length 384;

Best Local Similarity 100.0%; Pred. No. 1.2e-204; Indels 0; Gaps 0;
Matches 384; Conservative 0; Mismatches 20;

OY 1 MRETVLVLCLLEGEFAFDMDQYHIVWDEDFYFGAKWQHEVTARGGSEFOLATOP 60
DB 1 mrvtlvllclllegfafaftdwdqyhlwvdeidyfgakwqhevtarggsefqltgd 60
OY 61 GANSFYRDKLFIKPTLADNINPOTGAPRGDPMNGVLDVAMTGACTNNDNGCYRT 120
DB 61 gansfyrdkflfkptladdninpogtgaprgdpmngvldvamtgactnndngcyrt 120
OY 121 GAGSDTPMARSARVTFQYKSFTHGRVYVNAKMPVGDMLPAMTMMPEMDVYGGMPISGE 180
DB 121 gagsdtpmarsarvtfqyksfthgrvynakmpvgdmlpamtmmpemdvyggmpisge 180
OY 181 IDIIETIGNRDEKNTGEGEFLGQKMGSTMMGPGMDNRMTSLPKHSDMNYGDNFHT 240
DB 181 idiietigrndekntgegeflgqkmgstmmgpgmdnrmtslpkhsdmnygdnfht 240
OY 241 FWPDMSPNGLRFFVDDENQALDVPYPLIDANFMYVDFWEGKRWLPQYENDNPWAGTIN 300
DB 241 fwfdmspnglrffvddengalldvpyplidanpfmyvdfwegkrwlpqyendnpwagtn 300
OY 301 LAPFOQNFHILNVAVGTRNGFIPDCINRGSPALQKPKNSNDWYNDAMRKFFDAGNW 360
DB 301 lapfoqnfhilnvaavgtrngfipdcinrgspalqkpknsndwyndamrkffidargnw 360
OY 361 KWTWDEGDNNAQVYIRYRKRN 384
DB 361 kwtwdegdnnaqviryrykrn 384

RESULT 2

AAW29457 standard; Protein; 303 AA.

AAW29457;

14-APR-1998 (first entry)

Oerskovia xanthineolytica beta-1,3-glucanase.

Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;
fungal cell wall; intracellular product; purification; protoplast.

OS Oerskovia xanthineolytica LG109 (DSM 10297).

XX Key Location/Qualifiers

FT Peptide 1..52

FT Protein /label= sig_peptide

FT Protein 53..303

FT Protein /label= Mat_protein

PN MO9739114-A1.

PD 23-OCT-1997.

PF 14-APR-1997; 97MO-DK00160.

PR 23-AUG-1996; 96DK-0000885.

PR 12-APR-1996; 96DK-0000427.

PA (NOVO) NOVO-NORDISK AS.

PI Diers I, Ferrer P, Halkier T, Hedegard L;

DR MPI: 1997-526451/48.

DR N-PSDB: AAT89157.

XX New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia

XX xanthineolytica, used particularly for the lysis of microbial cells

XX for obtaining desirable products

XX Example 2; Page 42-43; 64pp; English.

XX This sequence comprises the polypeptide precursor of a novel

XX Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase

XX (BG) activity. Its amino acid sequence was deduced from an

XX isolated genomic DNA sequence (see AAT89157). Claimed DNA constructs

XX that encode the novel BG (see also AAW29455), a mannose binding

XX domain (see AAW29458) or a full-length enzyme, i.e. BG with mannose

XX binding domain (see AAW29456), can be used to produce recombinant BG

XX polypeptides, with or without a mannose binding domain, in fungal

XX or bacterial host cells. BG polypeptides are used for the

XX degradation or modification of beta-glucan containing material,

XX especially for the gentle lysis of microbial cell walls, thereby

XX enabling recovery of desirable intracellular products with a

XX reduced amount of contaminants. They can also be used for the

XX production of e.g. pigments, colourants, flavourants, yeast

XX extracts, pharmaceuticals, food or feed compositions, and to

XX prepare protoplasts for use in fusion, transformation and cloning

XX studies.

XX Sequence 303 AA;

Query Match 15.6%; Score 350; DB 18; Length 303;

Best Local Similarity 28.5%; Pred. No. 3.1e-25;

Matches 103; Conservative 30; Mismatches 96; Indels 132; Gaps 12;

OY 26 IWQDEDFYFDG-----AKMOHEVTATGGGSEFOLYTODGANSFVRDGLFIKPTLAD 80
DB 57 lhwdefdgaagsapnrvanvhetgahgwnaellqnylastransal-dgq----- 105
OY 81 NINPOTGAPFETDEMYGVLDVWAMGACTFTDNNGCYRTGACGIDIPPMASARVTFQY 140
DB 106 -----gnlvita-----fregdgsy-----tsamltgqky 131
OY 141 SFTHGRVYVNAKMPVGMVIMPLIMLPEMDVYGGMPRSGEIDIIETIGNRPFKNTGSEFL 200
DB 132 qpygyrleatrlqprgqivpafwlmglggsifpgrtwpssgeidienyvfephrvhg---- 187
OY 201 GICMGSTMMHGPMDNRWYLTSLPKHSDMNYGDNFHTFWMSPGRLFFVDDENQA 260
DB 188 -----tvh-gpysygs-gltgmyqbpqwsfadthtflavdkkpgelwfdv--gqg 236
OY 261 LLDVYPLIDANFMYVDFWEGKRWLPQYENDNPWAGTINLAPDONFHTLNVAVGTTN 320

Db	227	fhurverasvgnaw-----	-----vfqpflllnvavgqw	268
Oy	321	GFIDPGCINRGSDPALQKFWNSGDMYNDA	RRKFFDARGNWKWTWDESDNNMAQDY	IRV 380
Db	269	p9ypdg-----tttqp-----	-----gqmkvdvyrv	289
Oy	381	y 381		
Db	290	y 290		
	RESULT 3			
ID	AAW29456	standard; Protein: 435 AA.		
AC	AAW29456;			
XX	14-APR-1998	(first entry)		
DT				
	Oerskovia xanthineolytica beta-1,3-glucanase.			
XX				
XX	Beta-1,3-glucanase; lytic enzyme; yeast; beta glucan degradation;			
KM	fungal cell wall; Intracellular product; purification; protoplast.			
XX				
OS	Oerskovia xanthineolytica LG109 (DSM 10297).			
XX				
FH	Key	Location/Qualifiers		
FT	Peptide	1..52		
FT		/label- Sig-peptide		
FT	Protein	53..435		
FT		/label- Mat-protein		
FT	Domain	304..435		
XX		/label- Mannose-binding_domain		
PN	W09739114-A1.			
XX				
PD	23-OCT-1997.			
XX				
PF	14-APR-1997;	97WO-DK00160.		
XX				
PR	23-AUG-1996;	96DK-0000885.		
PR	12-APR-1996;	96DK-0000427.		
XX				
PA	(NOVO) NOVO-NORDISK AS.			
XX				
PI	Diers I, Ferrer P, Halkier T, Hedegaard L;			
XX				
XX	WPI: 1997-526451/48.			
XX	N-PSDB: AAT89156.			
XX				
XX	New isolated beta-1,3-glucanase enzyme - obtained from Oerskovia			
XX	xanthineolytica, used particularly for the lysis of microbial cells			
XX	for obtaining desirable products			
XX				
XX	Example 2; Page 39-40; 64pp; English.			
CC				
CC	This sequence comprises the polypeptide precursor of a novel			
CC	Oerskovia xanthineolytica enzyme that exhibits beta-1,3-glucanase			
CC	(BG) activity and which includes a mannose binding domain (MBD).			
CC	Its amino acid sequence was deduced from an isolated genomic DNA			
CC	sequence (see AAT89156). Claimed DNA constructs that encode the			
CC	novel BG lacking a MBD (see AAW29455 and AAW29457), a MBD (see			
CC	AAW29458), or the full-length enzyme can be used to produce recombinant			
CC	BG polypeptides, with or without a mannose binding domain, in fungal			
CC	or bacterial host cells. BG polypeptides are used for the			
CC	degradation or modification of beta-glucan containing material,			
CC	especially for the gentle lysis of microbial cell walls, thereby			
CC	improving recovery of desirable intracellular products with a			
CC	reduced amount of contaminants. They can also be used for the			
CC	production of e.g. pigments, colourants, flavourants, yeast			
CC	extracts, pharmaceuticals, food or feed compositions, and to			
CC	prepare protoplasts for use in fusion, transformation and cloning			
CC	studies.			

[illegible]


```

Db 397 aaggtief-rdgsitgsg---vtkpw-----rdarkasvhwfrhmsdwfrwsqg--- 443
OY 371 NAMQVDYIRV 380
   ::||::|
Db 444 -slvdfivkv 452

RESULT 6
AAR89137
ID AAR89137 standard; Protein; 467 AA.
AC AAR89137;
XX
XX 22-AUG-1996 (first entry)
DE
DE Bombyx mori full length LPS-binding protein.
XX
XX lipopolysaccharide binding protein; Bombyx mori; haemolymph; primer;
XX Enterobacter cloacae; PCR; polymerase chain reaction; amplification;
XX probe; LPS complex; septic shock; infection; transgenic plant; vine;
XX tobacco; tomato; potato; fungal infection; fungus.
XX Bombyx mori.
XX
XX Key Location/Qualifiers
XX FH Peptide 1..12
XX FT /note- "signal peptide"
XX FT Peptide 13..467
XX FT /note- "mature protein"
XX FT Modified-site 182
XX FT /note- "putative N-glycosylation site"
XX PN FR271032-A1.
XX
XX PD 15-DEC-1995.
XX
XX PF 09-JUN-1994; 94FR-0007083.
XX
XX PR 09-JUN-1994; 94FR-0007083.
XX
XX PA (INSP ) INST PASTEUR.
XX
XX PI Brey PT, Lee W;
XX PI WPI; 1996-060094/07.
XX DR N-PSDB; AAT10279.
XX DR

New protein from Bombyx mori that binds bacterial
lipopoly-saccharide - esp. used to treat septic shock, also DNA
encoding it, for producing transgenic plant(s) resistant to fungal
attack
PT
PT Claim 5; Page 42-43; 53pp; French.
XX
XX PS
XX CC This is the amino acid sequence of the full length lipopolysaccharide
XX CC (LPS) binding protein from Bombyx mori. The protein was isolated from
XX CC the haemolymph of fifth stage B. mori larvae after injection with
XX CC Enterobacter cloacae strain 57-9. Partial amino acid sequence was used
XX CC to generate PCR primers (AAT10280-1). These amplified a fragment of the
XX CC gene used as a probe to obtain the full length gene by screening a cDNA
XX CC library derived from B. mori previously injected with heat-killed
XX CC E. cloacae. The protein or fragments of it, can be used to treat diseases
XX CC associated with LPS complexes e.g. septic shock, to remove LPS from
XX CC products used for injection and to protect transgenic plants e.g. vines,
XX CC tobacco, tomato or potato, against fungal infections.
XX
XX SQ Sequence 467 AA;

```

Query Match 14.0%; Score 314; DB 17; Length 467;
 Best Local Similarity 27.3%; Pred. No. 1.4e-21;
 Matches 101; Conservative 58; Mismatches 143; Indels 68; Gaps 17;

```

OY 27 WMODEFDYFDGAKNQ-HEVTATGGNSSEFQLYTODGANSFVADKLFKFTLLADINPQ 85
Db 147 lfeeqidsldenwvqlegypipihpeypfvsygrnllvstagnlhi-----naklq 199
OY 86 TGAP-FGTFDMNGVLDVWAMAGACTNTDNNCGYRTGAGDI-PPAMSAVRVTFQKYSPT 143
Db 200 qhmpgflddslysgfln---lfsgetss-aacclkgagadllpprvsgrl-tsfigaft 254
OY 144 HGRVYVHAKMPYGVGLWMLPDMWYGGNP-RSGEIDILETIGN-----RDFKNT 195
Db 255 ygtveltraklpqgdwllpeylllepfllkkyssmnyasgvklacargnaelysgpndysnt 314
OY 196 GGEFLGIGKMGSTMFGWGD-DNRXYWLSLEPKHSDDMWYGNFHTFMDSPNGLRFRV 254
Db 315 -----vlygspimdlecrenfistkrrtdgtswgdsfhtysvqcpdfialsv 362
OY 255 DDENQALDVPYPLIDAPMWVDFWEMGKPMPLPOYENDPMAGTNLAPFDONFHLNV 314
Db 363 dgeewarveaprdal-----pavcahaprhllyqagsqmapfdhfytllyg 408
OY 315 AVGTNGEIPDGCINRGSDPALQKPMNSGDMWTNDAMRK----FPDARGMKWTWDEGDN 370
Db 409 aaggtief-rdgsitgsg---vtkpw-----rdarkasvhwfrhmsdwfrwsqg--- 455
OY 371 NAMQVDYIRV 380
Db 456 -slvdfivkv 464

```

```

RESULT 7
AAR1599
ID AAR1599 standard; Protein; 261 AA.
AC AAR1599;
XX
XX DT 18-JUN-1991 (first entry)
XX
XX DE Beta-1,3-glucanase.
XX
XX KW Alkalophillic; heat resistant.
XX
XX OS Bacillus sp.
XX
XX PN JP03053883-A.
XX
XX PD 07-MAR-1991.
XX
XX PF 20-JUL-1989; 89JP-0185928.
XX
XX PR 20-JUL-1989; 89JP-0185928.
XX
XX PA (SHKJ ) SHINGIJUTSU KAIHATSU.
XX
XX DR WPI; 1991-113290/16.
XX DR N-PSDB; AAQ11293.
XX
XX PT Heat resistant beta-1,3-glucanase gene DNA - derived from
XX PT alkali-compatible Bacillus sp.
XX
XX PS Claim 7; Fig 2; 11pp; Japanese.
XX
XX CC The new enzyme is derived from alkalophillic Bacillus Sp. The
XX CC enzyme acts on beta-1,3-glucans to form glucose and laminaribiose.
XX CC It is stable over a wide pH range with opt. activity at weakly
XX CC acidic pH and has good heat resistance.
XX
XX SQ Sequence 261 AA;

```

Query Match 13.8%; Score 310; DB 12; Length 261;
 Best Local Similarity 26.9%; Pred. No. 1.6e-21;
 Matches 100; Conservative 35; Mismatches 83; Indels 154; Gaps 17;


```

XX 19-MAR-1998.
PD 04-SEP-1997; 97WO-0515532.
PF 11-SEP-1996; 96US-0712072.
PR (NYBL-) NEW YORK BLOOD CENTER INC.
XX Goldstein J, Hurst R, Leng L, Lenny L, Zhu A;
PI WPI: 1998-207405/18.
DR N-PSDB; AAV22722.
XX New isolated endo-beta-galactosidase - from Flavobacterium
PT keratolyticus, used particularly for de-antigenising human
XX erythrocytes bearing A1 antigen for blood transfusions.
XX Claim 6; Fig 2A-2B; 45pp; English.
XX The present sequence represents the Flavobacterium keratolyticus
XX endo-beta-galactosidase (designated ENDO-A) protein. The ENDO-A
XX protein can be used in conjugation with alpha-N-acetylgalactosaminidase
XX to remove externally and internally linked A antigen on group A1
XX erythrocytes. The resulting erythrocytes may be transfused into
XX individuals who would be otherwise unable to safely tolerate a
XX transfusion of type A1 blood. The ENDO-A protein can also be used to
XX degrade keratan sulphate (e.g. in the food industry), or in blood typing
XX reactions.
XX Sequence 422 AA;
SQ
Query Match 12.8%; Score 286.5; DB 19; Length 422;
Best Local Similarity 27.0%; Pred. No. 5.1e-19;
Matches 103; Conservative 33; Mismatches 100; Indels 147; Gaps 15;
QY 10 LIFGEGFAFTDMDOYHIVWODEFDY---FDGAKMOHEVTATGGN---SEFQLY-TODGA 62
DB 45 llnatvatld---yellwsefinsggsfdstkwysadrgtvaawkymtsipaysqds 101
QY 63 NSFFVROGKLFKPTLLADNINPQTGAPGCTDPMYNGVLDVWAMGACNTNNGCYRTGA 122
DB 102 nlvlt-----mhnayv----- 112
QY 123 AGDIPPMASARVTFQKYSFTHGKRVVYAKMPVGDMLPAILMPLPED-WVYGMPRSGEI 181
DB 113 agdpvayhaggyvskmofsmtygkvevraktfgtgrspatltmmpepataygypscgei 172
QY 182 DIETIGNRDPFKNTGGEFLGIQKMGSTMHMGPMGMDNRYMLTSLPKSHDDNNYGDNFTF 241
DB 173 dsmeihvneav-----myhtlhngsvlnaag---gstasksatynttd-ytly 216
QY 242 WFDSPSGARFEVVDENQALLDVPYPLIDANPMWVDFEWGKPMPLPOYENDNPMAGNTL 301
DB 217 tmlwspndltfyt---nnsf-----qylayrvggggtqg 247
QY 302 APFDQNFHFLTNAVAG-GTNGFIPDGCINRGDPALQKPMGNGDMYNDAMKRFEDARGNW 360
DB 248 wplvfpfylllmgggsggwpatainadl-----pfs----- 278
QY 361 KWTWDEGDNNAMQVDYIRYTK 382
DB 279 -----mqvdyvryvyk 288

```

```

XX (1-3)-beta-D-glucan sensitive factor.
DE (1-3)-beta-D-glucan sensitive factor; antifungal agent;
XX (1-3)-beta-D-glucan sensitive factor; antifungal agent;
KW mycosis diagnosis.
XX Limulus sp.
XX Key Location/Qualifiers
FH Peptide 1..19
FT Peptide /label- sig_peptide
XX W09501432-A.
XX 12-JAN-1995.
PD 29-JUN-1994; 94WO-JP01057.
PF 29-JUN-1993; 93JP-0184403.
PR (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX Iwanaga S, Muta T, Oda T, Seki N;
PI WPI: 1995-060996/08.
DR N-PSDB; AA081334.
XX DNA encoding a polypeptide comprising a tetrapeptide motif at
PT least once - which may be used as an antibacterial and
PT antifungal.
XX Claim 9; Pages 23-32; 51pp; Japanese.
PS AA081334 encodes AAR67915 a (1-3)-beta-D-glucan sensitive factor, it
XX has a high affinity for the (1-3)-beta-D-glucan found in fungal
XX cell walls. The protein is therefore useful for clinically
XX diagnosing mycosis, and as an antifungal agent for the removal
XX of fungi.
XX Sequence 673 AA;
SQ
Query Match 9.3%; Score 209; DB 16; Length 673;
Best Local Similarity 21.8%; Pred. No. 2.2e-11;
Matches 85; Conservative 49; Mismatches 106; Indels 150; Gaps 15;
QY 5 LVVL-CLIFGEGFA---FTDMDOYHIVWODEFDYFDGAKMOHEV---TATGGNSEFQLYT 58
DB 2 lvlccvvlhvgvalccshcpkxqlwvdsdefltngissweteimgnnglmgvnehlqyyr 61
QY 59 QDGANSFVRDGLFKPTLLADNINPQTGAPGCTDPMYNGVLDVWAMGACNTNNGCY 118
DB 62 re-naegegklvt-----takredydgfk 85
QY 119 RTGAAGDIPPMASARVTFQKYSFTHGKRVVYAKMPVGDMLPAILMPLPEDWVYGMPRS 178
DB 86 yt-----sarlkqfidskwygkkaakmaipstfryvwmfmsgdntnyrtpss 135
QY 179 GEIDTIFETIGNRDPFKNTGGEFLGIQKMGSTMHMGPMGMDNRYMLTSLP-----KHSDDN 233
DB 136 geidfie-----hntune-----kvrgtlnw-----stpdgahahhnrsh 172
QY 234 -YGDNFHTFWDSPNGARFEVVDENQALLDVPYPLIDANPMWVDFEWGKPMPLPOYEND 292
DB 173 tngldyhlhysvekmssikvfnng-----gyfev 202
QY 293 NPMAGCTMLAPDQNFHFLTNAVAGGTNGFIPDGCINRGDPALQKPMGNGDMYNDAMK 352
DB 203 klggvgngksafinkvfvllmaiggn-----wpgfdvadea--- 239
QY 353 FFDARGNKKWTWDEGDNNAMQVDYIRYTK 382
DB 240 -fpaK-----myldyvyryvyg 253

```

RESULT 11

AA67918 standard; Protein; 233 AA.

AA67918;

14-SEP-1995 (first entry)

(1-3)-beta-D-glucan sensitive factor glucanase domain (4-236).

(1-3)-beta-D-glucan sensitive factor glucanase domain (4-236);

antifungal agent; mycosis diagnosis.

Limulus sp.

W09501432-A.

12-JAN-1995.

29-JUN-1994; 94MO-JP01057.

29-JUN-1993; 93JP-0184403.

(SEK) SEIKAGAKU KOGYO CO LTD.

Iwanaga S, Muta T, Oda T, Seki N;

WPI; 1995-060996/08.

DNA encoding a polypeptide comprising a tetrapeptide motif at least once - which may be used as an antibacterial and antifungal.

Claim 5; Page 40; 51pp; Japanese.

AA67917-R67919 are active fragments of AA67915, a (1-3)-beta-D-glucan sensitive factor, they have a high affinity for the (1-3)-beta-D-glucan found in fungal cell walls. The proteins are therefore useful for clinically diagnosing mycosis, and as antifungal agents for the removal of fungi.

Sequence 233 AA;

Query Match 9.3%; Score 208; DB 16; Length 233;

Best Local Similarity 21.3%; Pred. No. 6.7e-12; Matches 79; Conservative 45; Mismatches 98; Indels 146; Gaps 13;

```

23 QYHIVMODEDFYDGAKEHEV--TATGGSEFOLYTODGANSFVRDGLFIKPTLLAD 80
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 2 kwqlwsdeltngiswefemngnglmgwnealqyire--naqveggklyl----- 52
QY 81 NINPQTGAPGTDFMYNGVLDVWAMYGACTWTDNNGCYRTGACADIPRMSARVTFQKY 140
Db 53 -----takredydgfkyt-----sariktqgdk 75
QY 141 SFTHGRVYVAAKPVGMILPRAIMLPEDMWYGGWPRSGEIDIIETIGNRKFNTGGFEL 200
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 76 swkygkkaakmalpsftgyvwmfwmsgdnltyvpsgsgeidle-----hrltne-- 127
QY 201 GIDKMGSTMHNGPCMDNRYWLTSLP-----KHSDDMN-YGDNFHTFWFDMSPNGLRFEV 254
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 128 ---kvrqtlm-----stpdgahhmrresntngldyhllysvewnsstlvkfv 172
QY 255 DDEQALLDVPYPLIDANPMWVWDFEWGKFWLPDYENDNPMAGSTNLAEPDQNFHTILNV 314
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 173 ngn-----gyfewkiggyvngkxkaefnkfvallnm 202
QY 315 AVGGTNGFIDGCIINRGDPLAQKPSNGDWYNDAMRKFFPARAGNWKWTWDEGDNNAMQ 374
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 203 a1g9gn-----wpgfdvadea-----fpak-----my 223

```

QY 375 VDYIRVYK 382

Db 224 ldyvrvyq 231

RESULT 12

AA92310 standard; Protein; 341 AA.

AA92310;

01-APR-1999 (first entry)

Streptomyces sp. SGTase protein.

SGTase; beta-1,3-glucanase.

Streptomyces sp.

Key Location/Qualifiers

FT Peptide 1..64 /label= signal_peptide

FT Protein 64..341 /note= "SGTase"

JP10337184-A.

22-DEC-1998.

09-JUN-1997; 97JP-0151321.

09-JUN-1997; 97JP-0151321.

(DNIN) DAINIPPON INK & CHEM INC.

WPI; 1999-113611/10.

N-PSDB; AAX01742.

New recombinant DNA - used for recombinant production of beta-1,3 glucanase

Disclosure; Page 7-9; 10pp; Japanese.

This sequence represents a novel Streptomyces sp. SGTase protein which has beta-1,3-glucanase activity. A large amount of SGTase can be prepared using the recombinant DNA.

Sequence 341 AA;

Query Match 7.5%; Score 168.5; DB 20; Length 341;

Best Local Similarity 24.0%; Pred. No. 6.2e-08; Matches 105; Conservative 30; Mismatches 108; Indels 195; Gaps 27;

```

QY 1 MRWTLVVL-CILRGEGFAFT-----DMDQYHIVMODEFD-----YFDGA 38
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 37 lrrllvalagalagag-alcitptlaasvpppsgvtq---viadddgpkgsqvdgt 92
QY 39 KWQHEVATG-----GENSEFOLYTODGAN-SFVRDGLFIKPTLLADNINPOTGAPF 90
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 93 dwry-atgtygpggsnngtgeletmtanpenvsldgnglrltp----- 136
QY 91 GTDFMYNGVLDVWAMYGACTWTDNNGCYRTG---AAGDIPRMSARVTFQKYSFTHGR 146
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 137 -----trdgagnwtsgrietariddfqpaggatlr----- 165
QY 147 VVYAAKPFV---GD---WLPRAIML--PEDWYGGWPRSGEIDIIETIGNRDKNTGG 197
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 166 --vearidvprntvdaakgyrpatfmligapyrgdywmwpravgeldime-----nt-- 213
QY 198 EFLGIDKMGSTMHNG--PGMDNRYWLTSLPRKHSDDMWYGDNFHTFWFDMWSPNG----LR 251
   ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

```

Db 214 --gsmutvatmbcgtppgpcnetaglggttcggtcagfhtymewdrdsveeir 271
QY 252 EFVDD-----ENQALLDVPYPLIDANPMWVDFEWGKRPMLPOYENDNPMAGTNLAPF 304
Db 272 fslldhtfhtvrenq-----vdatw-----snat-- 296
QY 305 DQNFHFLINAVAGTNGFIPDGCINRGDPAALQKFPWSNGDWYNDAMRKFFDARGNMWWTW 364
Db 297 dhgfvyllnvamgg--gf-pdaf---g9gp----- 320
QY 365 DDEG--DNNAQVYIRV 380
Db 321 -dagtqpgshmlvdyvqv 337

RESULT 13
AAB52463
ID AAB52463 standard; protein; 294 AA.
XX
AAB52463;
DT 23-FEB-2001 (first entry)
DE Mycobacterium tuberculosis secreted protein #28.
XX
KM Mycobacterium tuberculosis secreted protein; MTSP; vaccine.
XX
OS Mycobacterium tuberculosis.
XX
PN MO200066143-A1.
XX
PD 09-NOV-2000.
XX
PF 04-MAY-2000; 2000MO-DS12197.
XX
PR 04-MAY-1999; 99US-0132479.
XX
PR 04-MAY-1999; 99US-0132503.
XX
PA (PUBL-) PUBLIC HEALTH RES INST NEW YORK.
XX
PI Gennaro ML, Gomez MJ;
XX
DR WPI; 2001-007151/01.
XX
PT Novel Mycobacterium tuberculosis secreted polypeptides and
XX
PT polynucleotides useful in diagnosis, treatment and prophylaxis of
XX
PT tuberculosis.
XX
Claim 11; Fig 1; 60pp; English.
XX
CC The present invention relates to Mycobacterium tuberculosis secreted
XX
CC proteins (MTSP), where the polypeptide has M. tuberculosis specific
XX
CC antigenic and immunogenic properties. Compositions of the invention may
XX
CC be useful for diagnosing Mycobacterium tuberculosis infection and as a
XX
CC vaccine against M. tuberculosis infection.
XX
SQ Sequence 294 AA;

Query Match 6.2%; Score 138.5; DB 22; Length 294;
Best Local Similarity 21.6%; Pred. No. 3.6e-05;
Matches 80; Conservative 43; Mismatches 102; Indels 145; Gaps 19;

QY 26 IWQDFD-----YFDGAKMO---HEV---TATGGGNSFOLYDQGANSEFRDGLFIK 74
Db 56 llhdefddgagsvpskxqyvhnrptknpvgfdrrpffgqyrdrqrnvldg----- 110
QY 75 PTLADNINPQTGAPETDWTNGVLDVWAMATGACTDNNNGCYRTGAAGDIPPMASARV 134
Db 111 -----nsnlvlatregnyfgslv-----hglvr-----gslgtwearl 146
QY 135 RTQKXSFTHGRVVAHAKMVGWMLPAIMWLPEDWYWGMP-RSGELDIETIGNDRFK 193

```

```

Db 147 ----kfn-----lapgmwpaawlndd-----pgrsgeldliewyn----- 180
QY 194 NTGGEFLGIQMGSTMHMGPCWDNRKYLISLPKHSDDMNTGDNFHTWFDWSPNGLRPF 253
Db 181 -----gtwpsgtvhanp---gtafetcpigvdgsw-----hmrvtwmpsgmyfw 224
QY 254 VD--DENQALLDVPYPLIDANPMWVDFEWGKRPMLPOYENDNPMAGTNLAPFDQNFHFI 311
Db 225 ldyadglepyfsvpatgledlneprwplndp-----gkvtlp-----v 264
QY 312 LNVAVGTFNGFIPDGCINRGDPAALQKFPWSNGDWYNDAMRKFFDARGNMWWTWDEGDNN 371
Db 265 lnlavgsq-----g9dpa-----tgsypq----- 285
QY 372 AMQVYIRVY 381
Db 286 -mlvdwvrvf 294

RESULT 14
AAB88406
ID AAB88406 standard; protein; 292 AA.
XX
AAB88406;
DT 19-AUG-1996 (first entry)
DE Trichoderma harzianum endo-1,3(4)-beta-glucanase.
XX
DE Yeast; fungus; enzyme; endo-beta-glucanase; EC-3.2.1.6; hydrolase;
XX
KM endo-1,3(4)-beta-glucanase; beta-glucan degradation; hydrolysis;
XX
KM detergent; surfactant; fungicide; antifungal; cleaning.
XX
OS Trichoderma harzianum CBS 243.71.
XX
PN MO9531533-A1.
XX
PD 23-NOV-1995.
XX
PF 11-MAY-1995; 95MO-DK00188.
XX
PR 11-MAY-1994; 94DK-0000546.
XX
PA (NOVO) NOVO-NORDISK AS.
XX
PI Andersen LN, Breinholt J, Christgau S, Dalboe H;
XX
PI Kauppinen MS, Kofoed LV, Olsen HS;
XX
DR WPI; 1996-010920/01.
XX
DR N-PSDB; AAT09876.
XX
PT DNA encoding endo-beta-glucanase from Trichoderma harzianum -
XX
PT useful, e.g., in food processing, as antifungal agent, in cleaning
XX
PT compans., etc.
XX
PS Claim 1; Page 40; 57pp; English.
XX
CC This polypeptide is immunologically reactive with antibodies
XX
CC raised against a pure endo-1,3(4)-beta-glucanase from T. harzianum
XX
CC CBS 243.71. Using this sequence, the protein may be expressed
XX
CC recombinantly in transformed host cells, particularly Aspergillus
XX
CC oryzae or Aspergillus niger. Typical applications of the
XX
CC recombinantly produced protein are preparation of protoplasts
XX
CC and yeast extracts, brewing, wine and press-juice manufacture, in
XX
CC foods and feeds, as antifungal agents, etc.
XX
SQ Sequence 292 AA;

Query Match 5.9%; Score 131.5; DB 17; Length 292;
Best Local Similarity 22.3%; Pred. No. 0.00017;
Matches 86; Conservative 40; Mismatches 109; Indels 151; Gaps 23;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:52:04 ; Search time 19.14 Seconds
(without alignments)
413.098 Million cell updates/sec

Title: US-09-596-101A-3

Perfect score: 2240
Sequence: 1 MKRTLVVLCILFEGGFATD.....DDEGDNNAMQVDYIRYKRN 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	357	15.9	321	2	US-08-712-072C-3 Sequence 3, Appl 1
2	351.5	15.7	262	1	US-08-392-828C-37 Sequence 37, Appl 1
3	351.5	15.7	262	3	US-09-330-945-37 Sequence 37, Appl 1
4	344.5	15.4	276	2	US-08-712-072C-4 Sequence 4, Appl 1
5	307	13.7	306	2	US-08-824-707-2 Sequence 2, Appl 1
6	286.5	12.8	285	2	US-08-712-072C-5 Sequence 5, Appl 1
7	272.5	12.2	422	2	US-08-712-072C-2 Sequence 2, Appl 1
8	208	9.3	654	1	US-08-392-828C-2 Sequence 2, Appl 1
9	208	9.3	654	3	US-09-330-945-2 Sequence 2, Appl 1
10	129.5	5.8	292	2	US-08-737-526-4 Sequence 2, Appl 1
11	129.5	5.8	292	4	US-09-098-580-4 Sequence 4, Appl 1
12	112.5	5.0	738	1	US-07-985-458-3 Sequence 4, Appl 1
13	110	4.9	243	1	US-09-286-690-10 Sequence 3, Appl 1
14	106	4.7	238	3	US-09-286-690-7 Sequence 10, Appl 1
15	100	4.5	237	1	US-08-103-998-4 Sequence 7, Appl 1
16	99	4.4	279	3	US-09-286-690-9 Sequence 4, Appl 1
17	98.5	4.4	239	1	US-08-103-998-2 Sequence 9, Appl 1
18	92.5	4.1	312	4	US-09-216-295-21 Sequence 2, Appl 1
19	92.5	4.1	371	4	US-09-104-308-1 Sequence 21, Appl 1
20	90.5	4.0	242	3	US-09-286-690-11 Sequence 1, Appl 1
21	90.5	4.0	429	1	US-08-745-977-4 Sequence 11, Appl 1
22	90.5	4.0	429	3	US-09-040-699A-4 Sequence 4, Appl 1
23	90.5	4.0	552	3	US-09-120-365-5 Sequence 4, Appl 1
24	90.5	4.0	552	4	US-09-515-039-5 Sequence 5, Appl 1
25	89.5	4.0	491	1	US-08-206-176-4 Sequence 5, Appl 1
26	89.5	4.0	553	2	US-08-661-052-16 Sequence 16, Appl 1
27	89	4.0	750	6	US-08-712-072C-3 Patent No. 5457037

28	89	4.0	751	6	US-08-941-445A-17	Patent No. 5457037
29	89	4.0	822	3	US-08-941-445A-17	Sequence 17, Appl 1
30	87.5	3.9	1052	4	US-09-255-502-7	Sequence 7, Appl 1
31	87	3.9	635	2	US-08-797-366-5	Sequence 5, Appl 1
32	87	3.9	635	2	US-08-956-268-5	Sequence 5, Appl 1
33	87	3.9	854	2	US-08-928-692-17	Sequence 17, Appl 1
34	86	3.8	534	4	US-09-124-541-1	Sequence 1, Appl 1
35	85.5	3.8	242	3	US-08-451-409A-1	Sequence 8, Appl 1
36	85	3.8	215	1	US-08-451-409A-1	Sequence 2, Appl 1
37	84	3.8	485	4	US-08-749-391-2	Sequence 2, Appl 1
38	84	3.8	485	4	US-09-390-200-2	Sequence 2, Appl 1
39	84	3.8	560	3	US-08-814-052-6	Sequence 6, Appl 1
40	84	3.8	560	3	US-08-812-823-6	Sequence 6, Appl 1
41	84	3.8	799	3	US-08-941-445A-15	Sequence 15, Appl 1
42	83.5	3.7	289	2	US-08-580-545B-4	Sequence 4, Appl 1
43	83.5	3.7	289	4	US-09-262-653A-4	Sequence 4, Appl 1
44	83.5	3.7	289	4	US-08-867-484A-2	Sequence 2, Appl 1
45	83.5	3.7	313	3	US-08-508-761B-34	Sequence 34, Appl 1

ALIGNMENTS

RESULT 1
US-08-712-072C-3
Sequence 3, Application US/08/12072C
Patent No. 5925541
GENERAL INFORMATION:
APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,072C
FILING DATE: 11-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A.
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 63475/97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 321 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: el3b, Bacillus circulans
US-08-712-072C-3

```

Query Match Similarity 15.9%; Score 357; DB 2; Length 321;
Best Local Similarity 26.2%; Pred. No. 2.8e-28;
Matches 100; Conservative 46; Mismatches 86; Indels 150; Gaps 13;

OY 25 HITWODEF-----DYFDGAKMOHEVATGGGNSSEFQLYTODGANSFYRDKLFK 74
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 64 NLIWQDEFNETTLDTSKNMYFTGYLLNDNDPATWGMGNMELQHTNSTQNVVYQDDKLNK 123
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 75 PPLLADNINPOTGAPGVDGMTNGVLDWAMYGACTNFDNNGCYRTGAAGDIPAMSARV 134
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 124 --AMNDSKSPQDP-----NRYAQYS-----SGKI 145
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 135 RRFQKSTFHGVVYVYHAKRPVGDWMLPAPMLPEDMYVGGPRSGEIDIIETIGNRDFN 194
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 146 NKKDLTSLKGYRVDPRAKLPTGDGWPAALMLPKDSYVGTWMAASEIDMEARGLPQSV 205
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 195 TGGEFLGIQKMGSTWAMGHPMDNRYWTLSLPKHSDDN-----YGDNFHFEPDWSPN 248
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 206 SG-----THRGQWPNVQ-----SSGDIHFPPGQTFPANDYHYISVWEED 247
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 249 GIRFVDE-----NQALLDVPYPLIDANPWWVDVFWEMGKFMPLPOYENDNFWAGCTNLA 302
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 248 NIKWYVDGKFEFYKVNQ-----QWYSTAARNPR-----A 277
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 303 PRDQNFHFLINAVAGT--NGEIPDGCINRGDPALOKPMGSDMYNDAMKRFDPARGNW 360
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 278 PDEFYFLIMNLAVGNDGGRTP--NASDIPA-----308
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 361 KWTWDEGDNNAMQVDYIRYK 382
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 309 -----TMQVDYIRYK 319
   :::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 2
US-08-392-828C-37
: Sequence 37, Application US/08392828C
: Patent No. 5795962
: GENERAL INFORMATION:
: APPLICANT: IWANAGA, SADAOKI
: APPLICANT: MUTA, TATSUSHI
: APPLICANT: SEKI, NORIAKI
: APPLICANT: ODA, TOSHIO
: TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
: TITLE OF INVENTION: THEREOF
: NUMBER OF SEQUENCES: 39
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
: ADDRESSEE: THIBEAULT
: STREET: 53 STATE STREET
: CITY: BOSTON
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/392,828C
: FILING DATE: 28-FEB-1995
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: CAMPBELL, PAULA A
: REGISTRATION NUMBER: 32,503
: REFERENCE/DOCKET NUMBER: FJN-003
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 248-7000
: TELEFAX: (617) 248-7100
: INFORMATION FOR SEQ ID NO: 37:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 262 amino acids
: TYPE: amino acid

```

```

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Protein
LOCATION: 1..262
OTHER INFORMATION: /note= "BGI A1 SEQUENCE (FIGURE 2)"
US-08-392-B28C-37

Query Match      15.7%; Score 351.5; DB 1; Length 262;
Best Local Similarity 26.2%; Pred. No. 7,6e-28;
Matches 100; Conservative 44; Mismatches 89; Indels 149; Gaps 14.

QY 25 HIWODEPD-YFDGAKMOHEY-----TATGGNSEFOLYTODGANSFVRDGKFLTK 74
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 4 NLIWDERNGTTLDTSKNWYETGYLLNDPATWGMGNALQHYTSTQVWVYQDDGKLTK 63
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 75 PTLADININPOTGAPFGIDFMTNGVLVDWAMTGACTINDNCCYRTGAAGDIPPMASAV 134
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 64 A-----NDSKSFQDP-----NRYAQS-----SGKI 86
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 135 RFEQSYSTGGHYVYNAKAPVGDMLPAIYMLPEDMVYGGMPRSGEIDIIETIGNRDFKN 194
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 87 NTKDLSTAKGYRVPRAKLPFGSDGYWPAIYMLPKSYICTMAASGEIDYMEARGRLPGSV 146
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 195 TCGEFLGIQKMGSTMHMGPGMDNRYWLTSLPKHSDWN-----YGDNFHTFWDWSPN 248
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 147 SG-----TIHFGQMPVNO-----SSGGDYHFPBGQTFANDYHYSVWEED 188
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 249 GLRFVYDE-----NQALLDVPYFLIDANPWWVDFWEMGKRWLPQYENDNFWAGSTNLA 302
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 189 NIKWVYDEKFEYKYVNO-----QWYSTAAPNPN-----A 218
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 303 PEDQNFHILYNAVAGT--NGEIPDGCINRGDPALQKPMNSGDWYNDAMRKFDFARGNW 360
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 219 PDEDFYILMLNLAVGNDGGRTP-----NASDIPA-----A 249
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 361 KWTWDEGDNNMAQVDYIRVYK 382
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB 250 -----TMQVDYVRVYK 260)
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 3
US-09-330-945-37
: Sequence 37, Application US/09330945
: Patent No. 6077946
: GENERAL INFORMATION:
: APPLICANT: IWANAGA, SADDAKI
: APPLICANT: MOTA, TATSUSHI
: APPLICANT: SEKI, NORIAKI
: APPLICANT: ODA, TOSHIO
: TITLE OF INVENTION: DNA ENCODING HORSESHOE CRAB
: TITLE OF INVENTION: AMEBOCYTE LYSATE FACTOR G SUBUNIT A
: NUMBER OF SEQUENCES: 39
: CORRESPONDENCE ADDRESS:
: ADDRESS: PATENT ADMINISTRATOR, TESTA, HURWITZ &
: ADDRESSEE: THIBEAULT, LLP
: STREET: 125 HIGH STREET
: CITY: BOSTON
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/330,945
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:

```


APPLICANT: Hedegaard, Lisbeth
APPLICANT: Halckier, Torben
APPLICANT: Aasenjo, Juan
APPLICANT: Sava, Demitris
TITLE OF INVENTION: No. 591688el enzyme with beta-1,3-glucanase activity
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 591688el No. 591688elisk of No. 591688th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,707
FILING DATE: 14-April-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4290, 204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 306 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-824-707-2

Query Match 13.7%; Score 307; DB 2; Length 306;
Best Local Similarity 26.2%; Pred. No. 3.1e-23;
Matches 95; Conservative 36; Mismatches 97; Indels 134; Gaps 13;

QY 26 IWNODEFDYDGA-----WQHEVTATGGGSEFOLYTQDANSYVRGKLFIRPTLLAD 80
DB 68 IANSDEFGAAGSAPNDVNMHETGAGGMAELQNTTSSVNSAL-DGQ----- 116
QY 81 NINPOTGAPGTDEMNGVLDVMAVYGACTDNNNGCYRTGAAGDIPPMASARYTFQKY 140
DB 117 -----GNLVITL-----QESDGSY-----TSARLTQGNV 142
141 SFHGRVYVHAKMPVGMILPAILMLPBDWVYGGPMSGEIDITETGNRDKFTGSEFL 200
DB 143 QPQFGRLEARLQIPRGOGISAFMMVGNALPDTPPTSGEIDIMENVGNAAHEVHG----- 198
QY 201 GLOKMGSTMHMGPGMD-DNRWLTSLPKHSDWYGNFHTFEMFDMNGSLREFVDDENO 259
DB 199 -----TVH-GPEYSGDN-GINGTYOHPOGMSFADFFHFGIDWYTGCEITWLDV--GQ 246
QY 260 ALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNPMAGTNLAPDQNFHFLINAVAGGT 319
DB 247 EYHVTYATADVANGM-----VFDQFFFLINVAIGG- 277
QY 320 NGFIPDGCINRGDPALQKPSNSGDWYNDAMRKFFDARGNKKWTWDEGDNNAQVDYIR 379
DB 278 -----QMPGNPDATTFPP-----QQMKVDYVR 299

RESULT 6
US-08-712-072C-5

Sequence 5, Application US/08712072C
Patent No. 5925541
GENERAL INFORMATION:
APPLICANT: Jack Goldstein, Alex Zhu and Lin Leng
TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amster, Rothstein & Ebenstein
STREET: 90 Park Avenue
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10016
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH 1.44 MB STORAGE DISKETTE
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/712,072C
FILING DATE: 11-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Bogosian, Elizabeth A.
REGISTRATION NUMBER: 39,911
REFERENCE/DOCKET NUMBER: 63475/97
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-5995
TELEFAX: (212) 286-0854 or 286-0082
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 285 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
AMBI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
ORGANISM: Flavobacterium keratolyticus
US-08-712-072C-5

Query Match 12.8%; Score 286.5; DB 2; Length 285;
Best Local Similarity 27.0%; Pred. No. 3.4e-21;
Matches 103; Conservative 32; Mismatches 100; Indels 147; Gaps 15;

QY 10 LIFGGAFTAFDQDQHYIWNODEFDY---FDGAKWQHEVTATGGN---SEFOLY-TODGA 62
DB 8 LLNATVTATTD---YELMSDFENSSGGFDSYKWSYADRGTYAANKKYSLSLPAVSAQDS 64
QY 63 NSFVBDGKLFKPTLLADNINQDGPAPGTDFMYNGVLDVMAVYGACTDNNNGCYRTGA 122
DB 65 NLVLR-----MDNAVY----- 75
QY 123 AGDIPPMASARYTFQKYSFHGRVYVHAKMPVGMILPAILMLPBD-WVYGGPMSGEI 181
DB 76 AGDPYAVHAGVYKSKGTSMTYGVKVEYRAKFTQGRGSPWALIMBPATAYGWPSCGEI 135
QY 182 DIIFETGNRDEKNTGGEFLGLOKMGSTMHMGPGMDNRYWLTSLPKHSDWYGNFHTF 241
DB 136 DSMEHVNNESV-----MYHTIHNSYVTANG--GSTASKSKSTYVTTD-YINLY 179
QY 242 WEDMSPNGLRFEVDENQALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDNPMAGTNL 301
DB 180 TWINSFPNDIRFVY---NNSL-----QYTVARVSGGGTQ 210
QY 302 APFQDNFHFILNAVVG-GTINGFIPDGCINRGDPALQKPSNSGDWYNDAMRKFFDARGNW 360

Db 211 WPFVPPYLLINQAGAGMPGATNADL-----PFS----- 241
QY 361 KMTWDEGDNNAQVYRYRK 382
Db 242 -----MOVDYRYRK 251

RESULT 7
US-08-712-072C-2
; Sequence 2, Application US/08712072C
; Patent No. 5925541
; GENERAL INFORMATION:
; APPLICANT: Jack Goldstein, Alex Zhu and Ian Leng
; TITLE OF INVENTION: ENDO-BETA-GALACTOSIDASE
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amster, Rothstein & Edenstein
; STREET: 90 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/712.072C
; FILING DATE: 11-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Bogosian, Elizabeth A.
; REGISTRATION NUMBER: 39,911
; REFERENCE/DOCKET NUMBER: 63475/97
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 286-0854 or 286-0082
; TELEFAX: TWX 710-581-4766
; TELEX: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 422 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-712-072C-2

Query Match 12.2%; Score 272.5; DB 2; Length 422;
Best Local Similarity 27.2%; Pred. No. 1.6e-19;
Matches 105; Conservative 29; Mismatches 97; Indels 155; Gaps 17;

QY 10 LLEFGFAFTMDQYHIWODEPDY---FDGAKQHEVTATGSGSEPOLYTDGANSFV 66
Db 45 LLMATYATFD---YELWSDSEFNSSGGFDSTKMSYADRGVAMNK--YMTSDGSNLVL 99
QY 67 RDGKLFKPTLLADININPOTGAPFCTDFMYNGVLVWVWAMYGACTTNDNGCYRTGAAGDI 126
Db 100 R-----MDNAV-----AGD- 108

QY 127 PPMSARRTTOKYSFTIGRVYNAK-----MPV---GDMILPAIWMLEPD-WYTGGR 177
Db 109 FVAHAGGVKMSKESMTYKVEVRAKFTGVSLPAYAGRGSPAIIMMPEPATAYAGWPS 168

QY 178 SGEIDIIETIGNRDKTNGEFLGIQKMGSTMHGPGMDNRYWLTSLPKHSDDMYNGDN 237
Db 169 CGEIDSMENVNNEV-----MYHTIHNSVTNANG---GSTASKSATYNTTD- 212
QY 238 FHTFWDSPNGLRFREVDENQALDVPYPLIDANFMYDFEMGKRPWLPOENDNPWAG 297
Db 213 YNLTYMTWSPNDIRFYV---NNSL-----QYTYARVSG 243

QY 298 GTNLAPFDONFEHILNAVVG-GTNGFIPDGCINRGDPALOKPWSNGDYNAMRKFFDA 356
Db 244 GTQMPFDVPPYLLINQAGAGMPGATNADL-----PFS----- 278

QY 357 RGNMKTWDEGDNNAQVYRYRK 382
Db 279 -----MOVDYRYRK 288

RESULT 8
US-08-392-828C-2
; Sequence 2, Application US/08392828C
; Patent No. 5795962
; GENERAL INFORMATION:
; APPLICANT: IWANAGA, SADAKI
; APPLICANT: MUTA, TATSUSHI
; APPLICANT: SEKI, NORIHI
; APPLICANT: ODA, TOSHIO
; TITLE OF INVENTION: NOVEL POLYPEPTIDE AND DNA ENCODING
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, TESTA, HURWITZ &
; STREET: 53 STATE STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/392.828C
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: CAMPBELL, PAULA A.
; REGISTRATION NUMBER: 32,503
; REFERENCE/DOCKET NUMBER: FJM-033
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 654 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-392-828C-2

Query Match 9.3%; Score 208; DB 1; Length 654;
Best Local Similarity 21.5%; Pred. No. 1e-12;
Matches 79; Conservative 45; Mismatches 98; Indels 146; Gaps 13;

QY 23 OYHIWODEFDYFDGAKQHEV--TATGGNSEFOLYTDGANSFYRDGKLFKPTLLAD 80
Db 5 KQOLWSDSEFTNGISDSMDFEMGNGNGNNELOQYRRE--NAOVEGGLTV----- 55

QY 81 NINPOTGAPFCTDFMYNGVLVWVWAMYGACTTNDNGCYRTGAAGDIPAMSAARVRFQY 140
Db 56 -----TAKREDYDGFYK-----SARLKTQFDK 78


```

Query Match 5.8%; Score 129.5; DB 2; Length 292;
Best Local Similarity 22.0%; Pred. No. 3,1e-05;
Matches 85; Conservative 40; Mismatches 110; Indels 151; Gaps 22

QY 24 YHIVODEDFYDGA-----KMQHEVTATGGNSFQLYTQDGANSFVRDCKLFIRKPTLL 78
Db 28 FNLVWTFDFPAGNGGTSPNQNNW-NITNLNLNVAEDETYSSTPAANVLSSGS-----TL- 80
QY 79 ADNINPQIGAPPGCTDPMYNGVLDVWALMYGACTYTDNNNGCIRTGAAGDIPRMSARVATFQ 138
Db 81 -QLVPMRDSKGT-----STFGGWT-----SERLES-- 105
QY 139 KYSFV-HGRVY-VHAKMPVG-----DLMLPAIMWLPEDWVY--GCMPSRGEDIDRET 186
Db 106 KTFEFPAGKVTIRLEAIRFGSNAQANKGIVAFWMLGSLGPGSGSWRNCISGIDIMET 165
QY 187 IGRNDFKNTGGEPL-----GIQKMGSTMHWP---GMDNRVWLTSLPKHSDDMNVG 235
Db 166 V---GGAATAGHDLHCDYVPGGICNGENGI-GGPVNIAVNNDMHAWVEIDRTPSSW--- 218
QY 236 DNFHFEMFQMSGNGLEFEFDDENQALLDVYPLIDANPMWVDEFWEMCKPMLPYENDNPNW 295
Db 219 ---QSETLWISLDGTTFYFQITGS-----RGNGGCVW 246
QY 236 AGGTNLAPEDQNFHFLTNVAVGGTNGFIDPGCINRGDPALQKFWNSGWDYNDAMRKEFD 355
Db 247 ---NNIA--HSPLEFILTNAVVG-----NMPGNP-----NSATL----- 275
QY 356 ARGNKKWTYDDEGDNNANQVDYIRY 381
Db 276 -----DQYGSMEVGYAQY 290

RESULT 11
US-09-098-580-4
; Sequence 4, Application US/09098580
; Patent No. 6140096
; GENERAL INFORMATION:
; APPLICANT: Kofoed, Lene Venke
; APPLICANT: Andersen, Lene No. 6140096b0e
; APPLICANT: Kaupplinen, Markus Sakari
; APPLICANT: Christgau, Stephan
; TITLE OF INVENTION: An Enzyme With Endo-1,3(4)-B-Glucanase
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61400960 No. 6140096disk of No. 6140096th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/098, 580
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/737, 526
; FILING DATE: 08-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Valela, Gregg A
; REGISTRATION NUMBER: 35,127
; REFERENCE/DOCKET NUMBER: 4174.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655

```

```

: TELE:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 292 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-09-098-580-4

Query Match      5.8%; Score 129.5; DB 4; Length 292;
Best Local Similarity 22.0%; Pred. No. 3,1e-05;
Matches 85; Conservative 40; Mismatches 110; Incls 151; Gaps 22.

OY  24 YHIVODEFDYFDGA-----KQHEVTAATGGGSEFQLYTODGANSFYVDGKLFYKPTLL 78
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  28 FVLVSTDFAGNGCGSPNNMNW-NITGTNLNNAEQEYSSSTANVVLGGGS-----TL- 80
      :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY  79 ADNINPQGAPRGCDPMFNGLVDVAMAGACNTDNNCGYRFGAAGDIPPAASAVRFQ 138
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  81 -QLVPMRDSKGT-----STEGWT-----SGRES-- 105
      :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY  139 KYSEF--HGRVY-VYAKKPVG-----DWLPAITMLPEDEVY--GGWPRSGEIDILET 186
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  106 KYTFPRAGKQYTRLEAARFSGNAAQANQKIGIPARFMLGDSLROGGSPNNGEIDIMET 165
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY  167 IGRNDFKNTGGEFL-----GIQKSGSTMWGP---GWDNRVYLSLPHSDQWNG 235
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  166 V---DGAQTATGHTLCHDYVPGGICNEGNGI--GGPVNIANVNDMHWVRVIDRTPSSW--- 218
      |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY  236 DNEHFHWDMSENGRLRFVYDDENQALLDVPYPLIDANPMWVDFWMEKRFWLFOYENDNF 295
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  219 ---QSEYTWLSIDGTIFYFOYIGS-----RIGQGW 246
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY  296 AGGTMLAPFDQNFHFLINVAAGTNGFLPDGGINSGDPALQKPMNSGDMYNDAMKRF 355
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  247 ---NNIA--HSPLEFLINVAAG-----NMNGN-----NSATL----- 275
      ||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

OY  356 ARGNNKWTWDEGDNNANQVDYIRVY 381
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB  276 -----DGYGSMEVGYVAY 290
      :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

RESULT 12
US-07-985-458-3
: Sequence 3, Application US/07985458
: Patent No. 5344777
: GENERAL INFORMATION:
: APPLICANT: Tamaki, Toshim;
: APPLICANT: Takemura, Hiroshi;
: APPLICANT: Takama, Kenji;
: APPLICANT: Fukaya, Masahiro;
: APPLICANT: Okumura, Hajime and
: APPLICANT: Kawamura, Yoshiya
: TITLE OF INVENTION: Structural Gene of Membrane-Bound
: TITLE OF INVENTION: Alcohol Dehydrogenase Complex, Plasmid
: TITLE OF INVENTION: Containing The Same And Transformed Acetic Acid
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fishauf, Holtz, Goodman & Woodward, P.C.
: STREET: 600 Third Avenue
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10016-2088
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 0.72 mb
: COMPUTER: IBM PC compatible (NEC PC-9801 ES)
: OPERATING SYSTEM: MS DOS
: SOFTWARE: ASCII form
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/985,458

```

1 FILING DATE: 19921203
2 CLASSIFICATION: 435
3 PRIOR APPLICATION DATA:
4 APPLICATION NUMBER: US 07/658,221
5 FILING DATE: 20-FEB-1991
6 APPLICATION NUMBER: 73440/1990
7 FILING DATE: 26-MAR-1990
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Goodman, Herbert
10 REGISTRATION NUMBER: 17081
11 REFERENCE/DOCKET NUMBER: 910134/MG
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: (212)972-1400
14 TELEFAX: (212)370-1622
15 TELEX: 236268
16 INFORMATION FOR SEQ ID NO: 3:
17 SEQUENCE CHARACTERISTICS:
18 LENGTH: 738 amino acids
19 TYPE: amino acid
20 TOPOLOGY: unknown
21 MOLECULE TYPE: protein
22 FEATURE:
23 NAME/KEY: MATURE PEPTIDE
24 LOCATION: 36 to 738
25 IDENTIFICATION METHOD: N-terminal sequences of the
26 IDENTIFICATION METHOD: purified protein having a molecular weight of about
27 ORIGINAL SOURCE:
28 ORGANISM: Acetobacter altoaceticus
29 STRAIN: MH-24
30 PUBLICATION INFORMATION:
31 AUTHORS: Tamaki, Toshimi;
32 AUTHORS: Fukaya, Masahiro;
33 AUTHORS: Takemura, Hiroshi;
34 AUTHORS: Tayama, Kenji;
35 AUTHORS: Okumura, Hajime;
36 AUTHORS: Kawamura, Yoshiya;
37 AUTHORS: Nishiyama, Makoto;
38 AUTHORS: Horinouchi, Sueharu and
39 AUTHORS: Beppu, Teruhiko
40 TITLE: Cloning and Sequencing of the Gene Cluster
41 TITLE: Encoding Two Subunits of Membrane-Bound
42 TITLE: Alcohol Dehydrogenase from Acetobacter
43 TITLE: polyoxogenes
44 JOURNAL: Biochimica et Biophysica Acta.
45 VOLUME: 1088
46 PAGES: 292-300
47 DATE: 1991
48 07-985-458-3

```

Query Match Similarity      5.0%; Score 112.5; DB 1; Length 738;
Best Local Similarity      20.4%; Pred. NO. 0.0062;
Matches 78; Conservative   37; Mismatches 113; Indels 155; Gaps 21.

QY    48 GGGNSEFQLYTODGANSFYR-----DGLKLFKPTLLADININPQTAAPEGT-D-FMYN 97
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    212 GNGGSEF-----GARGFYSAFDAETGKDWMFFIYP-----NPKNPDDAASDSVLNN 258
                                     ~~~~~~

QY    98 GYLDMVAMYGACTNDNNCC-----YRTGAAGD----- 125
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db    259 KAYQTWSPTGAATRGCGGTVWDSIYYDEVAIDLVLGVGSGSPMNKYKSEKGDLFLFG 318

QY    126 ----TPPMASAVNRFEQ-----KYSTHER-----VVYRAK----- 152
       ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db    319 SIYALKPETGEYVMHFOETPDOWMETSQQIMTLTDLPIINGETRHHIVARRKNNGFFYIID 378

QY    153 MPVGDMIPAIIMLPEDWYGGW-----PRSEIDIIETIGNRD--FKNTGGEFLIQ-K- 204
       ||| : : : ||| | | | | | | | | | | | | | | | | | | | | | |
Db    379 AKTGEEI-----SGSNYYVMMASGLDPKTG-----RIYHPDALYITLGKEWISIPED 427

QY    205 MG-----SYMHWGPBGDDNNRYITLSLP-----KHSDNNVYGDNFHTFWFD 244
       : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```

```

Db      428  LGGHNFAMARSP-----KGLVYIPRAQOVPFLYTNQVGGFTPHPDSDMNLGLDMNKVIP 482
Oy      245  WSPNGLREFVDDENQALIDVPYPLIDANPMKWVDFWE---NGKPMWLPQYENDNPPAGSTNL 3011
Db      483  DSPSQAQAFVVDLK-----GWIVAMDPQKQAEAW--RVDHKGPWNGGIIA 5255
Oy      302  APDPQNFHFLN---VAVGGTNG 321
Db      526  TGGDILFQGIANGGEFHAYDATING 548

RESULT 13
US-09-286-690-10
; Sequence 10, Application US/09286690
; Patent No. 6103511
; GENERAL INFORMATION:
; APPLICANT: Li, Xin-Liang
; APPLICANT: Ljungdahl, Lars G.
; APPLICANT: Chen, Huizhong
; TITLE OF INVENTION: Lichenase and Coding Sequences
; FILE REFERENCE: 55-96
; CURRENT APPLICATION NUMBER: US/09/286,690
; CURRENT FILING DATE: 1999-04-05
; EARLIER APPLICATION NUMBER: US 60/027,882
; EARLIER FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/US97/17811
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ. ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRP
; ORGANISM: Bacillus licheniformis
US-09-286-690-10

```

```

Query Match similarity      4.9%: Score 110; DB 3; Length 243;
Best Local Similarity     21.9%: Pred. No. 0.0023;
Matches      60; Conservative   42; Mismatches    86; Indels    84; Gaps    18;

QY      71 LEIKETLLADINIPOTGAPFGTDE-MYNGVLDMVAMYACTWTDDNNGCYRTGAAGDIPPA 129
        || : : | : | | - | | | : | : | : | : | :
DB      16 LFLSLSTFAASASMQTGSGFYEPFNYY--TGLWQKAGCYSIGNNFNC--TWRRANNVSMT 71
        : : : : : : : : : : : : : : : : : : : : :
QY      130 MSARVR-----TEOKY-----SFTHGKRVVNHACMPYGVWLMPA--IWMLEPDENVY 172
        : | : : : : : : : : : : : : : : : : : : :
DB      72 SLGEKRILSLTSPSYKKFPCGCENRSVOYTGGLYEVMNK-----PAKNVCIVSSFFTY 123
        : : : : : : : : : : : : : : : : : : : :
QY      173 GGMWRPG---EIDIIETIGNRDFKNITGGEPGLIOKMSTMHMGPGWGNDRKWLTSLPKH 228
        : | : : : : : : : : : | | | | | | : : : :
DB      124 TG-PTDGTPMDEIID-| | | | | | | | | | | | | | | | | | | |
        : : : : : : : : : : : : : : : : : : : :
QY      229 SDDNNYG---DNFHTFWEDMSPNGLRFPEVDE--NOALLDVYP--LIDANPW---WVD 277
        : : : : : : : : : : : : : : : : : : : :
DB      160 EKIYNLGCDANSYHTTFYAFDMQNPNSIKKYVDGQLKHRTITQTLPQTKIMNMNLWGAGVD 219
        : : : : : : : : : : : : : : : : : : : :
QY      278 FWEWGKPWLPOYLENDPWNAGSTNLAPFDONHFI 311
        || | | | | : | : | : | : | : | :
DB      220 -----EWLSGY-----NGVPLRLSRSLHW 238
        : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-09-286-690-7
: Sequence 7, Application US/09286690
: Patent No. 6103511
:
: GENERAL INFORMATION:
: APPLICANT: Li, Xin-Liang
: APPLICANT: Ljungdahl, Lars G.
: APPLICANT: Chen, Huizhong
: TITLE OF INVENTION: Lichenase and Coding Sequences
: FILE REFERENCE: 95-96
: CURRENT APPLICATION NUMBER: US/09/286,690
: CURRENT FILING DATE: 1999-04-05
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:55:27 ; Search time 25.22 Seconds
(without alignments)
1159.835 Million cell updates/sec

Title: US-09-596-101a-3

Sequence: 1 MRRTLVVLCILFGEGFAFTD.....DDEGDNNAMQVDYIRVYKRN 384

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Maximum number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 68:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	713	31.8	499	2 JC6141	beta 1,3-glucanase
2	369.5	16.5	877	2 JN0772	glucan endo-1,3-beta
3	353	15.8	1324	2 T18265	endo-1,3(4)-beta-g
4	351.5	15.7	682	2 J00420	beta-1,3-glucanase
5	344.5	15.4	286	2 S48201	licheninase (EC 3.
6	314	14.0	467	2 JC6150	gram-negative bact
7	284	12.7	642	2 B72428	lamellarinase - The
8	209	9.3	673	2 A49878	coagulation factor
9	179.5	8.0	466	2 T35164	probable secreted
10	148	6.6	364	2 T17584	probable beta-1,3-
11	138.5	6.2	294	2 D70525	probable beta-1-
12	132	5.9	419	2 T39920	probable glucanase
13	125.5	5.6	1144	2 A75132	hypothetical prote
14	125	5.6	857	2 B69798	hypothetical prote
15	116.5	5.2	878	2 A83748	endo-beta-N-acetyl
16	115	5.1	284	2 T06200	xyloglucan endo-1,
17	112.5	5.0	738	2 T06270	xyloglucan endo-1,
18	112.5	5.0	954	1 S20907	alcohol dehydrogen
19	111	5.0	282	2 A85354	endo-1,4-beta-xyla
20	110.5	4.9	856	2 T00349	hypothetical prote
21	110	4.9	243	1 S15388	avicelase III - As
22	109	4.9	781	2 A43866	licheninase (EC 3.
23	109	4.9	807	2 B82158	neuraminidase - Vi
24	108	4.8	282	2 T02354	neuraminidase VC17
25	108	4.8	636	2 T37843	xyloglucan endo-1,
26	108	4.8	750	2 S61143	probable beta-gluc
27	107.5	4.8	728	2 T48815	KRE6 protein - Yea
28	107.5	4.8	772	2 T02098	mixed-linked glucan
29	106	4.7	238	1 S19012	probable phosphatid
					licheninase (EC 3.

30	106	4.7	742	2 A49340	alcohol dehydrogen
31	106	4.7	782	1 G64157	probable organic s
32	105	4.7	802	2 A36910	xylinase, beta(1,3
33	104.5	4.7	239	1 A29091	licheninase (EC 3.
34	103.5	4.6	851	2 H84053	endo-beta-1,3-1,4
35	103	4.6	289	2 F71402	xyloglucan endo-1,
36	102.5	4.6	276	2 T40453	licheninase (EC 3.
37	102.5	4.6	334	1 S23498	licheninase (EC 3.
38	101.5	4.5	984	2 T44496	cellulose 1,4-beta
39	100.5	4.5	277	2 S71222	xyloglucan endo-1,
40	100.5	4.5	292	2 T06201	xyloglucan endo-1,
41	99	4.4	564	2 T49073	hypothetical prote
42	99	4.4	1087	2 S41797	cellulose 1,4-beta
43	98	4.4	1048	2 C86189	protein T25N20.11
44	98	4.4	1196	2 A29130	beta-amyase (EC 3
45	97.5	4.4	611	2 S60040	alpha-amyase (EC

ALIGNMENTS

RESULT 1	
JC6141	
beta 1,3-glucanase (EC 3.2.1.-) precursor - sea urchin (Strongylocentrotus purpuratus	
C/Species: Strongylocentrotus purpuratus (purple urchin)	
C/Date: 11-Apr-1997 #sequence-revision 09-May-1997 #text-change 21-Jul-2000	
C/Accession: JC6141; PC6037	
R/Bachman, E.S.; McClay, D.R.	
Proc. Natl. Acad. Sci. U.S.A. 93, 6808-6813, 1996	
A/Title: Molecular cloning of the first metazoan beta-1,3 glucanase from eggs of the	
A/Reference number: JC6141; MUID:96270625	
A/Accession: JC6141	
A/Molecule type: mRNA	
A/Residues: 1-499 <BAC>	
A/Cross-references: GB:U49711; NID:g1488256; PIDN:AAC47235.1; PID:g1488257	
A/Accession: PC6037	
A/Molecule type: protein	
A/Residues: 21-40;197-209;329-344 <BA2>	
A/Experimental source: egg	
C/Comment: This enzyme functions in several extracellular activities including autooca	
eferase enzymes in plants.	
C/Keywords: egg; glycosidase; hydrolase	
F/1-20/Domain: signal sequence #status predicted <SIG>	
F/21-499/Product: beta 1,3-glucanase #status predicted <MAT>	
Query Match	31.8% Score 713; DB 2; Length 499;
Best Local Similarity	42.7% Pred. No. 4.7e-50;
Matches 158; Conservative 48; Mismatches 88; Indels 76; Gaps 15;	
QY 26 IVMODEFDYEDGAKMOHEVITATGGNSEFQLYTODGANSFYVDGKLFKPTLLADINPQ 85	
DB 179 LIPEEDFSFLDIMEHEMTAGGGMFEYYTNNRSYSYRDKGLFKPTLTQKLC--- 235	
QY 86 TGAFFGDEMTNGYLDVWAMTGA--CTYTDNNGCYRTGAAG-IPPAASARTRTQKXSF 142	
DB 236 -----GECSSLSGTLDDGSSPANLCTGNANVYGCRTSSNNLMLPDIOSARITYESPF 290	
QY 143 THGRVYVHAKKPVGDMLPALMLPEDVYGGMPRSGEIDIIETIGNRDFKTTGEEFGCI 202	
DB 291 KYGLIEYAKLPTGDMLPALMLPKHNGYGEWPSGETLDVESGNADIKADGLSACV 350	
QY 203 QKMSTMHMGWDNDRYMLSLPKHSDNMYGDNFHTFWDMSGNGLRFVDEDSQALL 262	
DB 351 DQMGSTMHMGFWPLNGY-----PK-----THATKFYVDDE----- 381	
QY 263 DVPPPLIDANRWYDPMWGRKPLPQYEN-----DNPAAGTN-LAPDQNHFLINAV 316	
DB 382 ----LLLNWDP-ATGFWDLG-----EFENDAPGIDNPAYNPKLTPDQEEYLLINAV 431	
QY 317 GGTNGEFLPDGCLNRGD-----PALQKPSNGSDWYNDLMARKFEDARGMKMTWDEGDNN 371	
DB 432 GGVNIF-----GGLITTPA--KPSNDS--PTASKDFWSPNTWYPIWNN--GEBA 476	

[illegible]

A:Accession: F18265
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1324 <SCH>
 A:Cross-references: EMBL:X09732; NID:g1143428; PID:e19085; PIDN:CAA61884.1
 C:Genetics:
 A:Gene: l1cA
 C:Keywords: hydrolase; glycosidase

Query Match 15.8%; Score 353; DB 2; Length 1324;
 Best Local Similarity 27.2%; Pred. No. 1.9e-20;
 Matches 100; Conservative 51; Mismatches 93; Indels 124; Gaps 16;

QY 22 DQYHVMQDEPD--YFPGAKWQHEVTATGGNSEFOLYTODGANSFYRDKLFTKPLLLA 79
 DB 425 EEWRLVMSDEFNGSEINMANSYDDPTGRNNGVQSTQ--NAYTKDGALYEAR--- 479
 QY 80 DNINPQTAPEPTGPMYNGVLDMVAMYGACTNTENNNGCYRTGAAGDIPPMASARVTFQK 139
 DB 480 ---KEDIEPGEYVHT-----SSKLIRK 503
 QY 140 YSFTHGRVYVNAKMPVGDWLPALIMLPEDW-VYGMPSREIDITETGRDKKNTGGE 198
 DB 504 KSMVKGEKEIRAKKPOGGIIPALIMMPEDPEFGYTPKCEIDIMELTGEHP----- 556
 QY 199 FLGIQKMSSTWMMGPGMDNRXYMLTSLPKSHSDMNGDNPFTFWFMSPLNREFVDEN 258
 DB 557 ----DKITGTHFEHPKESQGYT-LP--EGQTFDEDFHVISIENPGEIRYIDGK- 607
 QY 259 QALLDVPYPLIDAPPMWYDFEWGKPMLPQYENDNPAGG--TNLAPFDONFHLINVAV 316
 DB 608 --LVHV-----ANDW-----YSRD-PYLADDTYVPAPFDQNFILINISV 644
 QY 317 GGTNGFTPDGCLNGGPDALQPKMSNGDWYTDARKKFFDARGNKKWTVDEEGDNNAAQVD 376
 DB 645 G-----GGMPG-----YPDETTVF-----PQMVVD 665
 QY 377 YIRVYKRN 384
 DB 666 YRVYQKD 673

RESULT 4
 J00420
 beta-1,3-glucanase Al precursor - *Bacillus circulans*
 C:Species: *Bacillus circulans*
 C:Date: 07-Sep-1990 #sequence.revision 07-Sep-1990 #text-change 15-Oct-1999
 C:Accession: J00420
 R:Yahata, N.; Watanabe, T.; Nakamura, Y.; Yamamoto, Y.; Kamimiyu, S.; Tanaka, H.
 Gene 86, 113-117, 1990
 A:Title: Structure of the gene encoding beta-1,3-glucanase Al of *Bacillus circulans* W
 A:Reference number: J00420; MUID:90185240
 A:Accession: J00420
 A:Molecule type: DNA
 A:Residues: 1-682 <YAH>
 A:Cross-references: GB:M34503; NID:g142972; PIDN:AAA22474.1; PID:g142973
 A:Experimental source: strain WL-12
 C:Comment: This enzyme, together with chitinase, is crucial for hydrolyzing yeast and
 C:Genetics:
 A:Gene: glcA
 F:1-38/Domain: signal sequence #status predicted <SIG>
 F:39-682/Product: beta-1,3-glucanase Al #status predicted <MAT>

QY 25 HIVQDEPD--YFDGAKWQHEV-----YATGGNSEFOLYTQDCANSFYRDKLFTK 74
 DB 424 NLIWMDEPNTGLTIDSKNNTETGYLLNDPATWGMGAELHGYTNSQNVYVQDGKLNK 483

Db 220 TPCGATCQAGFHTYRVEMDRSTSVETIRFSVDGNN-----FHTVRANQ--VD----- 265

QY 283 KPLPLQYENDNPMAGCTLAPFDQNEHFLVAVGCTGFLPDGCGINCGDPAALOKPMNSN 342

Db 266 -----ATTWNAATN-----HOYFILLVAVMG--GF-PDAF--GGGP----- 297

QY 343 GDWYNDAARKFEDARGNMKWTWDEG--DNNAOVDYIRV 380

Db 298 -----DGTGTPGHSMLVDYIVOV 314

RESULT 10

T17584

probable beta-1,3-glucanase (EC 3.2.1.-) - Chlorella virus PBCV-1

C:Species: Chlorella virus PBCV-1

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 12-Nov-1999

C:Accession: T17584

R:Graves, M.V.; Van Etten, J.L.

submitted to the EMBL Data Library, May 1999.

A:Reference number: Z18806

P:Accession: T17584

status: preliminary; translated from GB/EMBL/DBJ

olecule type: DNA

A:Residues: 1-364 <GRA>

A:Cross-references: EMBL:U42580; NID:G4028896; PIDN:ACG96462.1

A:Experimental source: specific host Chlorella strain NC64A

C:Genetics:

A:Note: A94L

C:Keywords: glycosidase; hydrolase

Query Match 6.6%; Score 148; DB 2; Length 364;

Best Local Similarity 22.2%; Pred. No. 0.00014;

Matches 85; Conservative 40; Mismatches 108; Indels 150; Gaps 20;

QY 24 YHIVMODEFD--YFDGAKW--OHEVTATGGGSEFOLYTDGANSFVRDGLFKPTLLA 79

Db 101 FEAMWDEFDEBEIDRTKMYIOPDIVYITGNROJHIYDSFTIEVSDTLHI-----IA 156

QY 80 DNIENQTAGPRTDPMYNGVLDVWAMYGACTNTDNNCGYRTGA-----AGDIPPMASA 132

Db 157 NNPG-----EVOYNE-----TSSNDQFYTSARINTKTGTHGMYGMEV 196

QY 133 RVRFPQKSTFHGRVYVNAKMPVGMPLPAILMLEDWYTGMPRSGEIDIIET----- 186

Db 197 NGTWNNTI-----RVEARLKAPRGVGAFAWMLPID--NSCFP--EIDIFETPYCERA 246

QY 187 -IG-----RRDFKNTGGEFLGIQKMGSTMHWGPMNDNRMYLMSLPKHSDDWNYGDNFHTF 241

Db 247 SMGTIVYAKDVR-----GISKHGTI-----TESTD-KFCDIYIVY 282

QY 242 WEDWSPNGLREFVDDENALLDVPPPLIDANPMWVDFWEMGRFWLPQYENDNPMAGTNL 301

Db 283 AVENADYIAFAAGDAE-----TPFVTV-----GK-----EIMACKDA 316

QY 302 -----APFQONFHTIIVANGTNGCFLPDGCGINRGDPAALOKPMNSGDWYNDAARKFEDAR 357

Db 317 NDTDAPYNNPFYIILNTSISGAMGIP-----LNDIFPA- 350

QY 358 GNMKWTWDEGDNNAOVDYIRV 380

Db 351 -----VLVDYIRV 359

RESULT 11

D70525

probable beta-1 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C:Accession: D70525

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:96293987

A:Accession: D70525

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-294 <COL>

A:Cross-references: GB:Z96800; GB:AL123456; NID:G3261800; PIDN:CAB09586.1; PID:e32167

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV0315

Query Match 6.2%; Score 138.5; DB 2; Length 294;

Best Local Similarity 21.6%; Pred. No. 0.00066;

Matches 80; Conservative 43; Mismatches 102; Indels 145; Gaps 19;

QY 26 IYMODEFD--YFDGAKW--HEV--TATGGNSEFOLYTDGANSFVRDGLFK 74

Db 56 LIFHDEFDPAGSVDPDSKMOVSNHTPIKNPVGFRPQFPGYRDSRONVFDG----- 110

QY 75 PRLADNINPOTGAPGRTDPMYNGVLDVWAMYGACTNTDNNCGYRTGAAGDIPPMASARY 134

Db 111 -----NSMLVLRATREGNRYFGALV-----HGLMR-----GGISTWEARI 146

QY 135 RFPQKSTFHGRVYVNAKMPVGMPLPAILMLEDWYTGMP-RSGEIDIIETIGNRDK 193

Db 147 ---KENC-----LAPGMPPAWMLSDND-----PERSSEIDILEYGN--- 180

QY 194 NTGGEFLGIQKMGSTMHWGPMNDNRMYLMSLPKHSDDWNYGDNFHTFWMDSNGLRFF 253

Db 181 -----GWPSGTYVNAH---DSTAFETCPGVDGGM-----HMKRTWMPSCGYFW 224

QY 254 VD--DENQALDVPYPLIDANPMWVDFWEMGRFWLPQYENDNPMAGTNLAPDQNEHTI 311

Db 225 LDVADGIEPEYFSPATGIEDLINEPIREMPEND-----GYKVP-----V 264

QY 312 LVAVAGCTGFLPDGCGINRGDPAALOKPMNSGDWYNDAARKFEDARGNMKWTWDEGDN 371

Db 265 LNLAVGSG-----GDPA-----TGSYPQE----- 285

QY 372 AMOVDYIRV 381

Db 286 -MLVDWYRVF 294

RESULT 12

T39920

probable glucanase precursor - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T39920

R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.

submitted to the EMBL Data Library, October 1997

A:Reference number: Z21891

A:Accession: T39920

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-419 <MCD>

A:Cross-references: EMBL:AL121794; PIDN:CAB57923.1; GSPDB:GNO0067

A:Experimental source: strain 972h-; cosmid c21B10

C:Genetics:

A:Gene: SPAC21B10.07

A:Map position: 2

Query Match 5.9%; Score 132; DB 2; Length 419;

Best Local Similarity 23.1%; Pred. No. 0.0034;

Matches 56; Conservative 31; Mismatches 81; Indels 74; Gaps 15;

QY 87 GAPFGTDFWNGVLDVWAMYGACTNTDNNCGYRTG--AAGDIPPMASA-----R 133

Db 137 GTFPFDGDEMNIT--PTGFGVGYLDNRSSAKGLISANSNVYMAADSKHNSGSRPS 194

QY 134 VRFPOKYSFPHGVV---VYAKKPVGDMPLPALMLPBDWYVGGWPSGSEIDIIETGNR 190

Db 195 IRLQSTQYEHGHFLIDLIH--LPYGGTWPAPFMTLGGD-----WPNGEIDIEGV--- 244

QY 191 DFKNTGGEFLGIQKMGSTMHNGPC--WDDNRMYLTLSPKSHDDP----- 232

Db 245 ---NVG-----TSNOVTLHTGDCCEMEDIKRYMTGALTNCNVADAPNSYNACGVENP 295

QY 233 ---NYGDNFH-----TFWDMSPNGLR--FFVDENQALLDVPYPLIDANPWWDFWENG 282

Db 296 SGPSYGEAFKRNKGGVFLDMRSEGRISWFF-----NRS--EIPEDITSGSPQPA---KWS 346

QY 283 KP 284

Db 347 EP 348

LT 13

A:5132

Hypothetical protein PAB1790 - Pyrococcus abyssal (strain Orsay)

C:Species: Pyrococcus abyssal

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C:Accession: A75132

R:Anonymous, Genoscope

Submitted to the EMBL Data Library, July 1999

A:Description: Pyrococcus abyssal genome sequence: Insights into archaeal chromosome stru

A:Reference number: A75001

A:Accession: A75132

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1144 <KAM>

A:Cross-references: GB:AJ248285; GB:AL096836; NID:95458067; PIDN:CAB49770.1; PID:ej151566

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB1790

Query Match 5.6%; Score 125.5; DB 2; Length 1144;
Best Local Similarity 18.4%; Pred. No. 0.038;
Matches 85; Conservative 49; Mismatches 132; Indels 197; Gaps 23;

QY 15 GFAFTMDQYHIWODEFDYF---DGAKMO-----HEVATGGGNS----- 52

Db 66 GLOGANLDKLYIAMDDEYLYIAIKNTNANKVAYGIGIDVGGEGTGESGDSWGKRN 125

QY 53 -----EFOLYTODGANSFVADGKLFKPTLLADNINPOTGAPFEGTDFMYNGVLDVWAMYG 107

Db 126 FTREIDYEVY-----FWMDDGKIM-----SDAFNMN-----CTDWEKYSISIEGEX- 167

QY 108 ACTNNDNNGCYRTGAAGDIPRAPSARVTFOKYSFTHGRVVAHAKMPVGMVLMPLMLP 167

Db 168 AYTGDSSSGIQ-----VLEVALP-----NSALGKGP 193

QY 168 E-----DMVYGGMPRSGEIDIIETIGNRDPKNGGEFLGIQKMGSTHMGCMDD-----N 218

Db 194 EKTAITTWI--AGEEGSSAVDTLPVDPISNINGE-----WTDADMLS 236

QY 219 RYMLTSLPK-----HSDMN-----YGNFHTFWMSPNGLRFFVDENQAL 261

Db 237 NFLEISITKVIYDGNLDMNKAELVAGVPSGIEBANLDRLVSDSNLYIAITNTNPAK 296

QY 262 LDVYPL-IDAN-----PW-----WVDEWEGKP 284

Db 297 YNPDYGAIVDNGSGIGTYPDAKKIYSGTYLPDIYIAEAQDALTWVGCKW--- 353

QY 285 WLPOYEEDNWPAGGTNLAPDQNHFLILNAVAGSTNGFIPDGCINRGDPRALQPMNSGD 344

Db 354 -----DGNEMITGTNINSVGEYAY-----IGDNGSI-----QTLEIKVPMW--- 389

QY 345 WYNDAMKRFDPARGNMTWTWDEBGNAM-----QVDYIRV 380

Db 390 ----ALGMPERKIALIIVAGNEDGNSAVDTLPVDPISIDYENI 428

RESULT 14

B69798

Hypothetical protein yeta - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: B69798

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizli, A.; Gal lech, J.; Harwood, C.R.; Henaut, A.; Hildebert, H.; Holstappel, S.; Hosono, S.; Hullo, M koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Laidino A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se akench, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Reference number: A69580; M01D:98044033

A:Accession: B69798

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-857 <KUN>

A:Cross-references: GB:Z99107; GB:AL009126; NID:92632866; PIDN:CAB12528.1; PID:g26330

A:Experimental source: strain 168

C:Genetics:

A:Gene: yeta

C:Superfamily: Bacillus subtilis hypothetical protein yeta

Query Match 5.6%; Score 125; DB 2; Length 857;
Best Local Similarity 19.8%; Pred. No. 0.029;
Matches 97; Conservative 53; Mismatches 162; Indels 178; Gaps 27;

QY 12 FGEFATMDQYHIWODEFDYFDGAK-----W----- 40

Db 289 FAHQTNVAMVDVKIV--QDSSDHYSLSKRTGKDYAVWMLHGRKAGVYAGKNGVAL 347

QY 41 -----QHEVYATAGGNSERQ--LYTODG-----ANSVRGKLEFI 73

Db 348 GLRYFEKYPALBITITGLAGSRPKMTIWLMPDDEANDLRHYTGNTFHVASAIEGDEMRS 407

QY 74 KPTLLADNINQGTGAPFG---TDFMYNGVLDVMA-----MYGACTWDNNGCYRT 120

Db 408 DPTGIA--NTNEISLACFSHMSDEVLNADKQAPPLIVEBPVY-----YES 455

QY 121 GAAG-----DIPRAPSARVTFOK-----YSFTHGRVVAHAKMPV----- 155

Db 456 KALGVWSIIDTSHPLKLELEQLDAAFLYKKEVQRWYGFHMGVDMHTYDPIRHMWR 515

QY 156 ---GDMPLPALMLPBDWYVGGWPSGSEIDIIETIG--NRPKNTG---GEFLGIQKMG 206

Db 516 YDLGYSAMONNELVPTTILMGAFFRSGREDFRMAEAATRTSTFDSHLEAYAGLSRH 575

QY 207 STMHNGPMDNRYMLSLPASHSDOWNYGDNFHTFWMSPNGLRFFVDENQALL----- 262

Db 576 NVVHMGCGCKEAKRISMAGL-----HFFYLYTG-----DDRGDLLETVK 615

QY 263 DVYPLIDANPWWDFWEMGRPWLPOYENDNP--WAGGTN--LAPD--QNFHFLILNAVAG 317

Db 616 DADVALVKTDPRA--FYEKGRH--PTHARTGPDMWAFCSNMLAEWERFENSEYTKIEFG 672

QY 318 GTNGFIPDGCINR-----GG-----DPALQKPNNGWYVNDAMKF--PDARGNW----- 360

Db 673 -----INCLKRLPLRLLSGPTREYDPATSMLHMGGIAGGYHMIATGAPQYMELA 725

QY 361 ----KMTWDD 366

Db 726 ELDDREMED 735

RESULT 15

A83748

endo-beta-N-acetylglucosaminidase BH0785 [Imported] - *Bacillus halodurans* (strain C-125)

C/Species: *Bacillus halodurans*

```
C:\Access\on: AR37A8
C:\date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
```

C:\ACCESS101: A03/40

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of t

A;Reference number: A83650; MUID:20263314

A;Accession: A83748

A;Status: preliminary

A;Molecule type: DNA

A;Kreslaues: 1-8/8 <SIO>

A: Experimental source:
a; cross-references: gb:

C:\Genetics:

ene: BH0785

Query Match	5.28;	Score 116.5;	DB 2;	Length 878;
-------------	-------	--------------	-------	-------------

Best Local Similarity 22.1%; Pred. No. 0.15;
Matches 89; Conservative 50; Mismatches 143; Indels 117

Matches 86; conservative 30; mismatches 143; indels 117; gaps 237

[illegible]

```

Db      14 VILVLPSSGFASSPESSYWPETLLDWSPEPD--PDARENRRSSIPLREYV----- 64
          ::|||  :|||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

OV 52 SEFOLYTODGANSFVRDGLFIKPTLLADNINPOT-GAPE--GTFMYNGVLDVWAMYGA 108

[illegible][illegible]

QY 109 CTNTDNNGCYRTGAGD---IPPAMSARVRTFQKYSFTHGRVVHAK---MPV-GDWLWP 161

```
Db 113 MV-----YWAGSAGEGITTPTS-----GDVIDAHRNGVPILGNVFFP 150
```

.....

Db 151 -PK-VYGGQE-----EWVDKMLVRGEDGSFPADKLLVAAEYYGFGDGFINQ 195

QY 220 YWLTSLPKHSDW-----NYGDNEHTFWD--WSPNGLRF--FVDDENQALLDVPY 266

105 ЕЩЕ СДЕЛАЮ МОЕ И ВИ СЕ И РАСЧУ И ТИ Д С И Т С И Н Д И Р И О Н В И Т Д Е Н А М Е О Л О С Н 35 Е

001 367 RTDANDETDETTISVDPRT DOWEIDNPIAACCENI APECONQUETI INVAIGCENCEI - - - 333

26 / PLIDANPMWVDFWEMGKPMLEPQYENDNPMAGGINLAFEDQNEHILNVAVGGINGETL --- 323

Search completed: October 11, 2001, 15:55:28
Job time: 286 sec

Job time: 286 sec

Db 424 NLIWDENGFTLDTSSKNWYETGYLLNDPATWGMGNELQHTYNTSTQNVYVQDCKLNIK 483

QY 75 PTLADININPQTGAAPFGDFMTNGLVDYWMATGACTNTDNCCTYTGAAAGDIPLPMASARY 134

Db 484 A-----NDKSKFPQDP-----NNRYAQS-----SGKI 506

QY 135 RTFOKSTFHGHVYVYHAKMPYGDYMLPAIWMLEPDMWYGGMPRSGEIDIIETIGNRDKN 194

Db 507 NTKDKLSLKYGHVDFRALPTGDDGYWPLMLPMLPKOSYIGTMAASELIDVMEARGSLPGSV 566

QY 195 TGGEFELGLOKMGSTHMGSPGMDNRTWLTSLPKHSDMN-----YGDNFTFWFDWSPN 248

Db 567 SG-----TIHFGGMVPMNQ-----SSGDDYHFPGCGFFANDYHYISVWNEED 608

QY 249 GLRFVYVDE-----NQLLDVLYPLIDANPMWVDFTWEGKFWLPDYENDNPAGCTNIA 302

Db 609 NIKWVVDKFFPKYKVTNO-----QWYSTAAENPN-----A 638

QY 303 PFDDQHFELTILWAVAGT--NGFIPDGCINRGDPAIQAQPMWSGDWYNDAMRKFPDARGNW 360

Db 639 PFDEFFYLLIMLAVAGNFDGKRP-----NASDIPA-----SGKI 669

QY 361 KWTWDEGDNNAMOVYDIYRYK 382

Db 670 -----TWQYDYVRYK 680

RESULT 2

GOB_RHOMR STANDARD; PRT; 286 AA.

AC P45798:

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-FEB-1996 (Rel. 33, Last annotation update)

DE BETA-GLUCANASE PRECURSOR (EC 3.2.1.73) (ENO-BETA-1,3-1,4 GLUCANASE) (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE) (LICHENASE).

GN BGAL.

OS Rhodothermus marinus.

OC Bacteria; CF3 group; Rhodothermus group; Rhodothermus.

OX NCBI_TaxID=29545;

RA Spilihaert R., Hreggvidsson G.O., Kristjansson J.K., Eggertsson G., Palsdottir A.;

RA *Cloning and sequencing of a Rhodothermus marinus gene, bgla, coding for a thermostable beta-glucanase and its expression in Escherichia coli.*;

RL Eur. J. Biochem. 224:923-930(1994).

CC -1- FUNCTION: SHOWS ACTIVITY ON LICHENAN, BETA-GLUCAN AND LAMINARIN BUT NOT ON CMC CELLULOSE OR XYLAN. THE ENZYME HAS A TEMPERATURE OPTIMUM OF 85 DEGREES CELSIUS AND A PH OPTIMUM OF 7.0.

CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLUCOSIDIC LINKAGES IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.

CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U04836; AAA60459.1; -.

DR HSSP; P23904; IAKK.

DR InterPro; IPR000757; -.

DR Pfam; PF00722; Glyco_hydro_16; 1.

DR PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.

KM Hydrolyase; Glycosidase; Signal.

FT SIGNAL 1 30 POTENTIAL.

FT CHAIN 31 286 BETA-GLUCANASE.

FT	ACT_SITE	158	158	NCBI
FT	ACT_SITE	163	163	PROTON DONOR (BY SIMILARITY).
SO	SEQUENCE	286 AA	33145 MM	7215C33624135191 CRR64;
Query Match	15.4%	Score 344.5	DB 1	Length 286;
Best Local Similarity	27.5%	Pred No. 2e-21	Matches 112	Conservative 37; Mismatches 96; Indels 163; Gaps
QY	1 MRWTVLVCLLFGGFATD-----	MDQYHWDEDFP----	DGAKWHEVATGG	50
DB	13 MRRAPFLSLVIGSGMSGRSDKAPHWE---	LWMSDEEDYSGLPDP	EKMVDYVGHWG	69
QY	51 NSEQLYTDGC-ANSFYRCKLFIKPTLLADNINPOTGAPSTDTMYNGVLDWANTGAC	109		
DB	70 NOELQYTRARIEARVGGVLLIEA-----	RHEPEGREY-----		105
QY	110 TINTDNNGCYRTGAAGDIPRMAKRVTFPKYSFTGGRVYVHAKMPVGDMLPAIWLPLED	169		
DB	106 -----	TSARLVTGRKASATYTGREFTIRLRDSGKSTYPAIMLPDR	145	
QY	170 WYVGG--WPSGEIDILLETIG-NRD-----	EKNTEGFEIIGKSGSTHWMGPGMDNRX	220	
DB	146 QYGSAYWPDNGEIDIMEHVGFPDVGHYVTKAYNHLLG--GQSGS-----	192		
QY	221 WLISLPHSDMKNGDNFHFPMDSPNGLRPFYDDE-----	NOALLDVPPLIDANPW	274	
DB	193 --INVPARFD-----	FHYTALTEPPEETIRMEVDSLYRREPNERLTD--	PEADWRHW	241
QY	275 WDFEWEMGKPMALPOYENDNPMAGCTMLAPDNFHFILVAVAGNGNGFIPODCINRGSDP	334		
DB	242 -----	FPDQPHILMIANGGANG-----	QGQVDP	267
QY	335 ALQPMNSGDMYNDAMKRFEDARGNMKWTDDEGDNNAQVDIRYK	382		
DB	268 -----	EAPPAQ-----	LWVDYVRYR	283
RESULT	3			
EXSH_RHIME	STANDARD	PRT	465 AA	
AC	03680;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	ENDO-1,3-1,4-BETA-GLYCANASE EXSH (EC 3.2.1.-)			(SUCCINOGLYCAN
DE	BIOSYNTHESIS PROTEIN EXSH).			
GN	EXSH.			
OS	Rhizobium meliloti (Sinorhizobium meliloti).			
OG	Plasmid pSymb (megaplasmid 2).			
OC	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;			
OC	Rhizobiaceae; Sinorhizobium.			
OX	NCBI_Taxid=382;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	York G.M., Walker G.C.;			
RT	"The Rhizobium meliloti exoX gene and prsD/prsE/exsH genes encode			
RT	components of independent degradative pathways which contribute to			
RT	production of low-molecular-weight succinoglycan."			
RL	Mol. Microbiol. 25:117-134(1997).			
RN	[2]			
RP	CHARACTERIZATION.			
RX	MEDLINE=98226741; PubMed=9560202;			
RA	York G.M., Walker G.C.;			
RT	"The Rhizobium meliloti exoX and exsH glycanases specifically			
RT	depolymetrize nascent succinoglycan chains."			
RL	Proc. Natl. Acad. Sci. U.S.A. 95:4912-4917(1998).			
CC	-I- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMM			
CC	SUCCINOGLYCAN. DYNAMICALLY REGULATE THE MOLECULAR WEIGHT			
CC	DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN			
CC	ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE			
CC	IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR			
CC	AGGREGATION STATE.			

CC -1- PATHWAY: EXOPOLYSACCHARIDE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: SECRETED. PROBABLY BY A TYPE-III SECRETION
 CC SYSTEM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: 089164; AAB64093.1; -
 CC InterPro: IPR000757; -
 CC InterPro: IPR001343; -
 CC Pfam: PF00722; Glyco_hydro_16; 1.
 CC PROSITE: PS00353; hemolysin_cabind; 1.
 CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; FALSE_NEG.
 CC PROSITE: PS00330; HEMOLYSIN_CALCIDIUM; FALSE_NEG.
 CC EXPOLYSACCHARIDE SYNTHESIS; GLYCOSIDASE; HYDROLASE; PLASMIID.
 CC DOMAIN 275 465 NUCLEOPHILE (BY SIMILARITY).
 CC ACT_SITE 349 349 CATALYTIC.
 CC ACT_SITE 354 354 PROTON DONOR (BY SIMILARITY).
 CC SEQUENCE 465 AA; 50286 MW; 6C8482366E9B8CA8 CRC64;
 SO
 Query Match 7.8%; Score 175.5; DB 1; Length 465;
 Best Local Similarity 21.6%; Pred. No. 2.8e-07;
 Matches 77; Conservative 40; Mismatches 102; Indels 137; Gaps 15;
 QY 27 VQDEPFYFGAKQHEVTATGGNSEFQLYTQDANSFVDRGKLFKPTLLADNINPQT 86
 DB 238 VMAKFI-----VMAPEKGTILSSNGPDQWYINPS-----VEPTA-----SVNPF 277
 QY 87 GADPGDFMTNGVLDVAMAYGA-CTNNDNGCYRTGAGADIPRMAVSARVFQKXSPFHG 145
 DB 278 -----VANGVLTITRAARSAEALQAEINQYDT-----SGMLTYSFPAQYIG 319
 QY 146 RYVYHAKMPYVDMLPAILMIPEDMYVGGWPRSGEIDIIETIGNRDNFKNGTEGFLGLOKM 205
 DB 320 YFEKRAMDPDQGVWPAFWLLPAD-----GSRP-----PELDVVEKRGD--SNT-----V 363
 QY 206 GSRMHNGPGDNDNRVWLTSLPKHSDDMNYGDNFTFWDSPNGIRFVVDENQALLDVP 265
 DB 364 IATVYH--SNETGSTSIENSVKAD-----ASGFHTYGVLTMEELIYWFDDAALARADTP 417
 QY 266 YPLIDANPMWVDWEWCKPMLPQYENDNPWAGTINLAFDQNFHILVAVGTFNGTIPD 325
 DB 418 SDNHD-----PMY-----MLVNLAVGGIAGTIPRD 441
 QY 326 GCINRGDPALQKFWNSGNDWYNDAMRRKFFDARGNMKMTWDEGDNNAQVDYIRYV 381
 DB 442 GL-----ADGSEKKIDYIKAY 457
 RESULT 4
 EGIC_RHIME STANDARD; PRT; 465 AA.
 AC 092302;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ENDO-1,3-1,4-BETA-GLUCANASE EGIC (EC 3.2.1.-) (SUCCINOGLYCAN
 DE BIOSYNTHESIS PROTEIN EGIC).
 GN EGIC.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC plasmid pSymA (megaplasmid 1).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 NCBI_TaxID=382;
 OX [1]
 RN RP SEQUENCE FROM N.A.

RC STRAIN-CXMI-105;
 RX MEDLINE=99413305; PubMed=10485295;
 RA Sharypova L.A., Yurel S.N., Keller M., Simarov B.V., Puehler A.,
 RA Becker A.;
 RT "The eft-482 locus of Sinorhizobium meliloti CXMI-105 that influences
 RT symbiotic effectiveness consists of three genes encoding an
 RT endoglucanase, a transcriptional regulator and an adenylate cyclase.";
 RL Mol. Gen. Genet. 261:1032-1044(1999)
 CC -1- FUNCTION: CLEAVES HIGH MOLECULAR WEIGHT SUCCINOGLYCAN TO YIELD LMW
 CC SUCCINOGLYCAN. DYNAMICALLY REGULATES THE MOLECULAR WEIGHT
 CC DISTRIBUTION OF SUCCINOGLYCAN BY CLEAVING NASCENT SUCCINOGLYCAN
 CC ONLY DURING A LIMITED PERIOD AFTER ITS SYNTHESIS, PERHAPS BEFORE
 CC IT UNDERGOES A TIME-DEPENDENT CHANGE IN ITS CONFORMATION OR
 CC AGGREGATION STATE (BY SIMILARITY).
 CC -1- PATHWAY: EXOPOLYSACCHARIDE BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: SECRETED. PROBABLY BY A TYPE-III SECRETION
 CC SYSTEM (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AJ225896; CAB38101.1; -
 CC InterPro: IPR000757; -
 CC InterPro: IPR001343; -
 CC Pfam: PF00722; Glyco_hydro_16; 1.
 CC Pfam: PF00353; hemolysin_cabind; 1.
 CC PRINTS: PR00313; CABINDGRPT.
 CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; FALSE_NEG.
 CC PROSITE: PS00330; HEMOLYSIN_CALCIDIUM; FALSE_NEG.
 CC EXPOLYSACCHARIDE SYNTHESIS; GLYCOSIDASE; HYDROLASE; PLASMIID.
 CC DOMAIN 275 465 NUCLEOPHILE (BY SIMILARITY).
 CC ACT_SITE 349 349 PROTON DONOR (BY SIMILARITY).
 CC ACT_SITE 354 354
 CC SEQUENCE 465 AA; 49602 MW; 1BC4A640685E9654 CRC64;
 SO
 Query Match 7.5%; Score 167; DB 1; Length 465;
 Best Local Similarity 21.4%; Pred. No. 1.4e-06;
 Matches 84; Conservative 38; Mismatches 113; Indels 158; Gaps 18;
 QY 16 FAFTMDQYHI-----VQDEPFYF---DGAK-----WOHEVTATGGG 50
 DB 196 FANTSAADOLHANOFSLADRSVLTPTFSDPTLTQSDSTSGVMPDKYMAPEKAGTLLG 255
 QY 51 NSEFQLYTQDANSFVDRGKLFKPTLLADNINPQTGA-PFGTDFMYNGVLDVAMAYGA- 108
 DB 256 NDELQWY-----VNPT-----YQTPASANPFS---VTGDLVLTITAKPASQ 292
 QY 109 CTNNDNGCYRTGAGADIPRMAVSARVFQKXSPFHGRVYVAKMPYVDMLPAILMPE 168
 DB 293 AIQAEFTNGYDYLT-----SGMLTYSFPAQYIGYFEMRADMPDQGANPAFWLLPAD 342
 QY 169 DMVYGGWPRSGEIDIIETIGNRDNFKNGTEGFLGLOKMSTMGWPGMDNDRVWLTSLPKH 228
 DB 343 D---GTWP-----PELDVVEKRGD--PNTVYATVHNSNETGQTS-----IASARV 384
 QY 229 SDDMNYGDNFTFWDSPNGIRFVVDENQALLDVPYPLIDANPMWVDWEWCKPMLPQ 288
 DB 385 TDT-----SGFHYGVLTMEELIYWFDDAALARADTPSDMHD-----PM 424
 QY 289 YENDPMWAGTINLAFDQNFHILVAVGTFNGTIPDGCINRGDPALQKFWNSGNDWYND 348
 DB 425 Y-----MLVNLAVGGIAGTIPRD-----GG----- 446
 QY 349 AMRKFEDARGNMKMTWDEGDNNAQVDYIRYV 381
 DB 447 -----AEKQVDYIKAY 457


```

RESULT 5
DHET_ACCEP STANDARD; PRT; 738 AA.
ID DHET_ACCEP
AC P28036;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ALCOHOL DEHYDROGENASE [ACCEPTOR] PRECURSOR (EC 1.1.99.8).
GN ADHA.
OS Acetobacter polyoxogenes.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Acetobacter.
OX NCBI_Taxid:439;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN-NB11028;
MEDLINE-91159482; PubMed-2001402;
Tamaoki T., Fukaya M., Takemura H., Tayama K., Okumura H.,
Kawamura Y., Nishiyama M., Horinouchi S., Beppu T.;
*Cloning and sequencing of the gene cluster encoding two subunits of
membrane-bound alcohol dehydrogenase from Acetobacter polyoxogenes.*;
RL Biochim. Biophys. Acta 1088:292-300(1991).
CC -1- FUNCTION: DEHYDRATION OF PRIMARY ALCOHOLS (EXCEPTION: METHANOL).
CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR -> ALDEHYDE +
    REDUCED ACCEPTOR.
CC -1- COFACTOR: POO AND HEME.
CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
    SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
    COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
    SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POO DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL, D00635; BAA00528.1; -
CC DR PIR, S14270; S14270.
CC DR HSP, P22619; 2MTA.
CC -----
CC InterPro: IPR000345; -
CC InterPro: IPR001479; -
CC InterPro: IPR002372; -
CC Pfam: PF01011; Bacterial_POO_6.
CC DR PROSITE, PS00363; BACTERIAL_POO_1; 1.
CC DR PROSITE, PS00364; BACTERIAL_POO_2; 1.
CC DR PROSITE, PS00190; CYTOCHROME_C; 1.
CC KM Oxidoreductase; POO; Heme; Periplasmic; Membrane; Signal.
CC FT SIGNAL 1 35
CC FT CHAIN 1 36 738 ALCOHOL DEHYDROGENASE [ACCEPTOR].
CC FT BINDING 650 650 HEME (COVALENT) (BY SIMILARITY).
CC FT BINDING 653 653 HEME (COVALENT) (BY SIMILARITY).
CC FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC FT METAL 654 654 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
CC SO SEQUENCE 738 AA; 80840 MW; 1E2B6ED7BCD92AF6 CRC64;

Query Match 5.0%; Score 112.5; DB 1; Length 738;
Best Local Similarity 20.4%; Pred. No. 0.072;
Matches 78; Conservative 37; Mismatches 113; Indels 155; Gaps 21;

```

```

OY 48 GAGNSEFOLYTDGANSFYR-----DGKLFKPTLLADININPOTGAPFGND-FMYN 97
DB 212 GNGSGSEF-----GAGPVSAPDAEKGKVDWRFYTP-----NRKNEDDASDSVLAM 238
OY 98 GVLVDWAMYGACTNTDNNCG-----YRTGAGD----- 125
DB 259 KAVQWSPFGAMTRGCGGGVWMDSIYDPAVLVYLGVGNGSPWNKRYRSEGGDNFLG 318

OY 126 -----IPPAMGARVTRFO-----KSTFTHR-----VYVHAK----- 152
DB 319 SIYALKRPGTGYVWHFETPMDDWFTSDQIMTLIDLPINGETRHYIYHARKNGFFYIID 378
OY 153 MPVGDMLPMAIWMLEPMVYGGW-----PRSGEIDITETGNRD--FKATGGEFLGID-K 204
DB 379 AKTGEFI-----SKRIVYVYMAASGLDPKIG-----RPIYNPDALYLTGKEWIGIFGD 427
OY 205 MG-----STMHWPGMDNDRRWLTSLP-----KHSDDMYGDNFHTFWEFD 244
DB 428 LGHNFAAMAFSP-----KTGLYIIPAOQVPELYTNQVGFTPHDPSWNLGLDMKVGIP 482
OY 245 WSPVGLTFFYDDENQALLDVPYPLIDANPMWVDWE---NGKPHLPQYENDNPPAGGTNL 301
DB 483 DSPPAKQAFVYKDLK-----GVIYAMDPOKQAEAW--RVDHGGPMNGGILA 525
OY 302 APFDONFETILN--VAVGTING 321
DB 526 TGGDLRFQGLANGEFHAYDAING 548

RESULT 6
XNLA_RUMFL STANDARD; PRT; 954 AA.
ID XNLA_RUMFL
AC P29126;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE BIFUNCTIONAL ENDO-1,4-BETA-XYLANASE XNLA PRECURSOR (EC 3.2.1.8).
GN XNLA.
OS Ruminococcus flavefaciens.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Ruminococcus.
OX NCBI_Taxid:1265;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-17;
RX MEDLINE-92261318; PubMed-1584021;
Zhang J.-X., Flint H.J.;
RT "A bifunctional xylanase encoded by the xynA gene of the rumen
RT cellulolytic bacterium Ruminococcus flavefaciens 17 comprises two
RT dissimilar domains linked by an asparagine/glutamine-rich sequence.";
RL Mol. Microbiol. 6:1013-1023(1992).
CC -1- FUNCTION: XYLANASE DOMAIN 1 RELEASES MORE XULO-OLIGOSACCHARIDES
    AND DOMAIN 2 MORE XULOSE.
CC -1- CATALYTIC ACTIVITY: ENDODIHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
    LINKAGES IN XYLANS.
CC -1- PATHWAY: XLAN DEGRADATION.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
    G (FAMILY 11 OF GLYCOSYL HYDROLASES).
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO CELLULASE FAMILY
    F (FAMILY 10 OF GLYCOSYL HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
    or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL, Z11127; CAA77476.1; -
CC DR PIR, S18043; S18043.
CC DR PIR, S20907; S20907.
CC DR HSP, P48793; 1XND.
CC DR InterPro: IPR001000; -
CC DR InterPro: IPR001137; -
CC Pfam: PF00331; Glyco_hydro_10; 1.
CC DR Pfam: PF00457; Glyco_hydro_11; 1.
CC DR PRINTS; PR00134; GLHYDRLASE10.
CC DR PRINTS; PR00911; GLHYDRLASE11.
CC DR PROSITE, PS00591; GLYCOSYL_HYDROL_F10; 1.

```


QY 271 ANPMWVDEWENG-----KPLDQ--YENDNPMAGCTNLAPD-----ONEHFI 311
 DB 301 SNTVYVSTARRPTDAQNGDRIRKIPWMPNGIFYSYDVASGNQAPLDYDQKERSEFQI- 359
 QY 312 LNVAVGTNGEIPDCINNGGDPALQKPMWSNGDWYNDARKKFFDAGCN 359
 DB 360 --AGMGSELYRRNTSLN-----SQDQMSNAKRIYDGAAN 394

RESULT 9
 KRE6_YEAST STANDARD: PRT: 720 AA.
 AC P32486;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE BETA-GLUCAN SYNTHESIS-ASSOCIATED PROTEIN KRE6 (KILLER TOXIN-RESISTANCE PROTEIN 6).
 GN KRE6 OR YPR159M.
 .cc Saccharomyces cerevisiae (Baker's yeast).
 .cc Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 .cc NCBL_TaxID=4932;
 .cc [1]
 .cc RP SEQUENCE FROM N.A.
 .cc RX MEDLINE=92107936; PubMed=1837148;
 .cc RA Roemer T., Bussey H.;
 .cc RT "Yeast beta-glucan synthesis: KRE6 encodes a predicted type II membrane protein required for glucan synthesis in vivo and for glucan synthase activity in vitro.";
 .cc RL Proc. Natl. Acad. Sci. U.S.A. 88:11295-11299(1991).
 .cc [2]
 .cc RP SEQUENCE FROM N.A.
 .cc RX MEDLINE=95176711; PubMed=7871892;
 .cc RA Roemer T.D., Fortin N., Bussey H.;
 .cc RT "DNA sequence analysis of a 10.4 kbp region on the right arm of yeast chromosome XVI positions GPH1 and SGV1 adjacent to KRE6, and identifies two novel tRNA genes.";
 .cc RL Yeast 10:1527-1530(1994).
 .cc [3]
 .cc RN CHARACTERIZATION.
 .cc RP MEDLINE=95014745; PubMed=7929594;
 .cc RA Roemer T., Paravicini G., Payton M.A., Bussey H.;
 .cc RT "Characterization of the yeast (1-->6)-beta-glucan biosynthetic components, Kre6p and Skn1p, and genetic interactions between the PKC1 pathway and extracellular matrix assembly.";
 .cc RL J. Cell Biol. 127:567-579(1994).
 .cc .- FUNCTION: INVOLVED IN THE SYNTHESIS OF (1->6)- AND (1->3)-BETA-D-GLUCAN POLYMERS OF THE YEAST CELL WALL. IN VIVO, IT IS REQUIRED FOR FULL ACTIVITY OF BETA-GLUCAN SYNTHASE IN VITRO. IT MAY BE A BETA-GLUCAN SYNTHASE, PART OF A MULTIPROTEIN GLUCAN SYNTHASE OR A MODULATOR.
 .cc .- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI SUBCOMPARTMENT.
 .cc .- PTM: PHOSPHORYLATED.
 .cc .- SIMILARITY: STRONG, TO SKN1.
 .cc .- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 .cc CC EMBL: M80657; AAA34726.1; .
 .cc DR EMBL: L33835; AAB59312.1; .
 .cc DR PIR: A41624; A41624.
 .cc DR SGG: S0006363; KRE6.
 .cc KW Glycoprotein; transmembrane; Cell wall; Signal-anchor; Phosphorylation.
 .cc FT DOMAIN 1 252 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 253 273 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT FT (POTENTIAL).
 FT DOMAIN 274 720
 FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 538 538 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 563 563 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 691 691 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 720 AA; 80178 MW; 958989EF7E9ACD69 CRC64;

Query Match 4.8%; Score 108; DB 1; Length 720;
 Best Local Similarity 19.0%; Pred. No. 0.16;
 Matches 82; Conservative 48; Mismatches 124; Indels 178; Gaps 23;

QY 36 DGAKMOHEVTAATGGNSEFQOLYTODGANSFVRDQKPIKFTLLADINPOTGAPFGTDM 95
 DB 322 DGSKW-----ELVFSDEFNA---EGRTFY-----DGDPPYWTAP--DVH 355
 QY 96 YNGVLDV-WAMTGACTNTD-----NNGCYRTGAAGDIPRAMSAVRFQKYSFT 143
 DB 356 YDATTQDLEWYSPDASTTNGTQLRLRDAPKKNHLY-----YRSGMLQSMNKVCFT 405
 QY 144 HGRVYVHAKMP-VG--DWLPAIWLPE-----DWYGGMPRS----- 178
 DB 406 QGALFISANLPLNPGRYSGPLGMLTNGNIGRPGLASTGQVWYSYSCDAGITPNQSSP 465
 QY 179 -----GEDIETITGNRDPKNTKGCEFLGIQMGST 208
 DB 466 DGISYLPQGLKSLCTCDVEDHPNQGVRGAPFELDVE--GETPTK-----IGVGLASQS 517
 QY 209 MHMGPGWMDNRWLTSLPRKHSDDMNNGDNFHTFPMSPGRLFFEDDEQALLDVPYPL 268
 DB 518 LQIAP-----FDIWMYPDYDFIYVFTTTMTATYAGGPPQOAVSAVSTL 562
 QY 269 IDANPMWVDEWENG--KPLDQYENDNPMWAGCTNLAPEDONFHLNVAVGSTNGFI-- 323
 DB 563 ---NVTWYEEGEGYGFQYKAYALEYND-----DDNGY--IRWVGDTPTFTYIH 605
 QY 324 ----PDGCTINNGGDPALQKPMWS-----NGCW-YNDARKKFFDAGCNKMTWDEGDN 370
 DB 606 AKALHPDG--NIGWRRISEPMISILNLGISNMWAYIDMOYIEFPV----- 649
 QY 371 NAMQVDYIRVYK 382
 DB 650 -VMSIDYVRIYQ 660

RESULT 10
 GUB_PAEPO STANDARD: PRT: 238 AA.
 AC P45797;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE BETA-GLUCANASE PRECURSOR (EC 3.2.1.73) (ENDO-BETA-1,3-1,4 GLUCANASE) (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHIDROLASE) (LICHENASE).
 GN GLOB.
 OS Paenibacillus polymyxa (Bacillus polymyxa).
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Paenibacillus.
 .cc NCBL_TaxID=1406;
 .cc [1]
 .cc RP SEQUENCE FROM N.A.
 .cc RX STRAIN=ATCC 842;
 .cc MEDLINE=92041687; PubMed=1938968;
 .cc RA Gosalbes M.J., Perez-Gonzalez J.A., Gonzalez R., Navarro A.;
 .cc RT "Two beta-glucanase genes are clustered in Bacillus polymyxa: molecular cloning, expression, and sequence analysis of genes encoding a xylanase and an endo-beta-(1,3)-(1,4)-glucanase.";
 .cc RL J. Bacteriol. 173:7705-7710(1991).
 .cc .- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLYCOSIDIC LINKAGES IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.

CC -1- MISCELLANEOUS: BETA-GLUCANASES OF BACILLUS HAVE A SUBSTRATE RANGE
CC SIMILAR TO LICHENASE OF GERMINATING BARLEY.
CC -1- SIMILARITY: BELONGS TO FAMILY 16 OF GLYCOSYL HYDROLASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X57094; CAA40379.1; -
CC HSP: P23904; 1CPM.
CC InterPro: IPR000757; -
CC Pfam: PF00722; Glyco_hydro.16; 1.
CC PRINTS: PR00737; GLHYDRLASE16.
CC PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
CC SIGNAL: 1 26
CC POTENTIAL.
CC CHAIN 27 238 BETA-GLUCANASE.
FT ACT_SITE 129 129 NOCLOPHILE (BY SIMILARITY).
FT ACT_SITE 133 133 PROTON DONOR (BY SIMILARITY).
FT DISULFID 56 85 BY SIMILARITY.
SQ SEQUENCE 238 AA; 26919 MW; C0CE7BAE5D40B8C CRC64;

Query Match 4.7%; Score 106; DB 1; Length 238;
Best Local Similarity 22.0%; Pred. No. 0.07;
Matches 62; Conservative 32; Mismatches 92; Indels 96; Gaps 16;

QY 25 HYWDEPFYDCAKMOHETATGGNSEFQLYTQDGAN-SYVRDGLFIKPLTLADNIN 83
DB 28 NFWF-PLPLSFNSSTWQ--KADGYSNGQWPCNCTWRANNFTNDGLKLTSSPANN-- 81
QY 84 PQTGAFTGDFMNGYLDVWAMTACTNTDNGCYRTGAAGDIPPMASARVTFQKSF- 142
DB 82 -----KFDG--GEYRSTNNY-YGLYEVSMKPAKNTGVSSFTYTG 121
QY 143 THGRVYVNAKMPYGDMLPRAIMLPEDWYGGMPRSGEIDIIETIGNRDRTNGEPLGI 202
DB 122 SHG-----TQW-----DEIDI-----EPLG- 136
QY 203 QKMSGMHMGDMNRWYLTSLPKHSDDMNYG--DNHTEFMDSMPLGRFVVD-- 256
DB 137 -KDTYVQF-----NTYNGVGGHEKITNLGFDASTSFHTAFDQPGITIKTYDGVLT 188
257 ENQALDVPPLIDANPMWVDFEWG---KPLPQYENDNP 294
DB 189 KHTATNIP-----STPGKIMMLMNGTGVDSMLGSGVNGANP 225

RESULT 11
DHET_ACEEU STANDARD; PRT; 739 AA.
AC 044002: 007952:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, last sequence update)
DT 15-JUL-1999 (Rel. 38, last annotation update)
DE ALCOHOL DEHYDROGENASE [ACCEPTOR] PRECURSOR (EC 1.1.99.8).
GN ADH.
OS Acetobacter europaeus.
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconacetobacter.
OX NCBI_TaxID=33995;
RX (1)
RP SEQUENCE FROM N.A.
RC STRAIN=DES11 / DSM 6160;
RA Thurner C.A.K.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: PRIMARY ALCOHOL + ACCEPTOR = ALDEHYDE +
CC REDUCED ACCEPTOR.
CC -1- COFACTOR: POQ AND HEME (BY SIMILARITY).

CC -1- SUBUNIT: HETEROTETRAMER (DEHYDROGENASE, CYTOCHROME, AND TWO
CC SMALLER UNKNOWN SUBUNITS) THAT FORMS THE ALCOHOL DEHYDROGENASE
CC COMPLEX (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC
CC SPACE (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL POQ DEHYDROGENASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X82894; CAA58066.1; -
CC EMBL: Y09480; CAA70688.1; -
CC InterPro: IPR000345; -
CC DR InterPro: IPR001479; -
CC DR InterPro: IPR002372; -
CC Pfam: PF01011; Bacterial_POQ; 6.
CC PROSITE: PS00363; BACTERIAL_POQ_1; 1.
CC DR PROSITE: PS00364; BACTERIAL_POQ_2; 1.
CC DR PROSITE: PS00190; CYTOCHROME_C; 1.
CC Oxidoreductase; POQ; Heme; Periplasmic; Membrane; Signal.
FT CHAIN 1 35
FT BRINDING 36 739 ALCOHOL DEHYDROGENASE [ACCEPTOR].
FT BRINDING 651 651 HEME (COVALENT) (BY SIMILARITY).
FT METAL 654 654 HEME (COVALENT) (BY SIMILARITY).
FT METAL 655 655 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 739 AA; 80944 MW; E681BB237ACB91F4 CRC64;

Query Match 4.7%; Score 106; DB 1; Length 739;
Best Local Similarity 20.6%; Pred. No. 0.25;
Matches 78; Conservative 37; Mismatches 118; Indels 146; Gaps 20;

QY 48 GGGSEFQLYTQDGANSEFVR-----DGKLIKPLTLADNINP----- 84
DB 212 GNGGSEF-----GAKFTVAFDAETGKVDWREFTAP-----NPKNEPHITASDSYLM 258
QY 85 -QTGAPPGT-----DFWYNGYLDVWAMTACTNTDNGCYRTGAAGD-- 125
DB 259 NKAQTMTSPGAMTRQGGGGVWDSIYDPAVDL-VYLGVNGSGPMWYKYSSEKGNLIF 317
QY 126 -----TPPMASARVTFQ-----KSFTHGRVYVNAKMPYGDMLPRAIMLP-- 167
DB 318 LGSIVALKPETGEYVWHFOETPMQMDFTSVQJMTLDLPINGETRHVIAHPKNGFEYI 377
QY 168 -----EDWYVGW-----PRSGEIDIIETIGNRD--FKNTGGEFLGIQ-KMG-- 206
DB 378 IDAKTGEFISGKNYVYVNMASGLDPKGTG-----RPYINPDALYLTGKEMWIGIEDDLAGH 432
QY 207 --STMHMGPGWDMNRWYLTSLP-----KHSDDWNYGDNHTEFMDSMSPN 248
DB 433 NFAAMAFSP-----KTGLVYVPAQGVPELYTNGVGFPHDPDSMNLGLDMNKVGIPODPE 487
QY 249 GLRFVVDENQALDVPPLIDANPMWVDFE---WGKPLPQYENDNPVAGGTILAFPE 305
DB 488 AKQAFVVDLK-----GWLAVMDPOKQAEAW--RVDHKGPMWNGILATGSD 530
QY 306 QNFHFIILN---VAVGTING 321
DB 531 LIFGLANGERHAYDATING 549

RESULT 12
OSTA_HAEIN STANDARD; PRT; 762 AA.
AC P44846;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DT 30-MAY-2000 (Rel. 39, last annotation update)

DE ORGANIC SOLVENT TOLERANCE PROTEIN HOMOLOG PRECURSOR.
 IMP OR OR STA OR H10730.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae.
 OC Haemophilus.
 OX NCBI_Taxid=727;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-RD / KM20 / ATCC 51907;
 RX MEDLINE=95350630; Pubmed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 Weidman J.F., Phillips R., Spriggs C.A., Hedblom E., Cotton M.D.,
 Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghegan N.S.M.,
 Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 Venter J.C.;
 RA "Whole-genome random sequencing and assembly of Haemophilus
 influenzae Rd".
 RT Science 269:496-512(1995).
 CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE (BY SIMILARITY).
 CC -1- SIMILARITY: STRONG, TO THE E. COLI HOMOLOG.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U3756; AAC22389.1; -
 DR TIGR: H10730; -
 DR Outer membrane; signal.
 KW SIGNAL 1 23
 FT CHAIN 24 782
 FT POTENTIAL.
 FT ORGANIC SOLVENT TOLERANCE PROTEIN
 FT HOMOLOG.
 SQ SEQUENCE 782 AA; 90084 MW; E73C8A5786B02D1B CRC64;

Query Match 4.7%; Score:106; DB 1; Length 782;
 Best Local Similarity 21.3%; Pred. No. 0.26; Indels 148; Gaps 26;
 Matches 83; Conservative 46; Mismatches 112;

DB 32 FDYFDGAKMOHEVYATGGNSEFQLYTODG--ANS--FV--RDGKLFKPTLLADNPQ 85
 106 FDYKD-----NQIMML-GKDAEFNLSHDGMLTSEYEFVGROGR-----GKADNIT-- 151
 86 TGAPFGTDFMYNGVLDVYAMTGACTNTDNGCYRTGANGADIPRANSA-RVRTFOKYSFT- 143
 152 -----LHNN-----TRYKMNATFTSLH-----GDVAMVADASEIRQYEEYAE 191
 144 --HGRVVVHA-----KAPVG-----DWLT-----PAIMML----- 166
 192 MHAHRAFIHGVPYVYTPPLQPLTGDRRSGLLISAGTSSODGLMYAOPITWINLAPNTDL 251
 167 ---PEDWYVGMPRSGEIDLETIGNRDKFKYTGGEFLGKQKSGTMHNGPGWDDNRMYLT 223
 252 TETFKYMSRGMQANGERRVLTSGE--GKVGAGEYLG-----KVRSSEVASDR----- 298
 224 SLPHASDMMYGDNPFHTFPMSPNGLRFYVDDENOALLDVYPLIDANPMKWDVFE-WG 282
 299 --KRHLFYWMNNSFLQNW-----RLNINVTYVSQKRYFNEDSDSYIG 338
 283 KP---WLPQYENDMPWAGSTMLAPEDONFHLILYAVAGSTNGF--IPDCINNGGPA--- 335
 339 KSTGICANQI-----ANATAYIQPNYNSLS-----AHQFQIPDDIVNIGRYRAVPQ 384
 336 -----LQKPMNSGDMYNDAMKRFPD 355
 DB 385 LDENVHKKYDLANGMLNFKLHSAQVAVR--FD 411

RESULT 13
 ID XYND_RUMFL STANDARD: PRT; 802 AA.
 AC 05317;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE XYLANASE/BETA-GLUCANASE PRECURSOR (INCLUDES: ENDO-1,4-BETA-XYLANASE
 DE (EC 3.2.1.8) (XYLANASE); (ENDO-BETA-1,3-1,4-GLUCANASE) (EC 3.2.1.73)
 DE (1,3-1,4-BETA-D-GLUCAN 4-GLUCANOHYDROLASE) (LICHENASE)).
 GN XYND.
 OS Ruminococcus flavefaciens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 OC Ruminococcus.
 OX NCBI_Taxid=1265;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-17;
 RX MEDLINE=93259938; Pubmed=8491715;
 RA Flint H.J., Martin J., McPherson C.A., Daniel A.S., Zhang J.-X.;
 RT A bifunctional enzyme, with separate xylanase and beta(1,3-1,4)-
 RT glucanase domains, encoded by the xynD gene of Ruminococcus
 RT flavefaciens".
 RL J. Bacteriol. 175:2943-2951(1993)
 CC -1- FUNCTION: CONTAINS TWO CATALYTIC DOMAINS WITH XYLANASE AND ENDO-
 CC BETA-1,3-1,4-GLUCANASE ACTIVITIES.
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
 CC LINKAGES IN XYLANS.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GLYCOSIDIC LINKAGES
 CC IN BETA-D-GLUCANS CONTAINING 1,3- AND 1,4-BONDS.
 CC -1- PATHWAY: XYLAN DEGRADATION.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO CELLULOSE FAMILY
 CC G (FAMILY 11 OF GLYCOSYL HYDROLASES).
 CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO FAMILY 16 OF
 CC GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: S61204; AAB26620.1; -
 DR HSSP: P23904; IATK.
 DR InterPro: IPR000757; -
 DR InterPro: IPR001137; -
 DR Pfam: PF02018; CBD_6; 1.
 DR Pfam: PF00457; Glyco_hydro_11; 1.
 DR Pfam: PF00722; Glyco_hydro_16; 1.
 DR PRINTS: PR00737; GLHYDRASE16.
 DR PRINTS: PR00911; GLHYDRASE11.
 DR PROSITE: PS00776; GLYCOSYL_HYDROL_F11_1; 1.
 DR PROSITE: PS00777; GLYCOSYL_HYDROL_F11_2; 1.
 DR PROSITE: PS01034; GLYCOSYL_HYDROL_F16; 1.
 DR Xylan degradation; Hydrolase; Glycosidase; signal;
 KM Multifunctional enzyme.
 FT SIGNAL 1 31
 FT CHAIN 32 802
 FT DOMAIN 32 244
 FT DOMAIN 245 523
 FT DOMAIN 524 555
 FT DOMAIN 556 802
 FT ACT_SITE 124 124
 FT ACT_SITE 226 226
 FT ACT_SITE 684 684
 FT ACT_SITE 524 529
 FT DOMAIN 532 543
 FT DOMAIN 546 553
 FT POLY-THR.
 FT POLY-THR.
 SQ SEQUENCE 802 AA; 89091 MW; 2880A689647284AF CRC64;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 11, 2001, 15:56:43 ; Search time 34.86 Seconds

(without alignments)
1457.405 Million cell updates/sec

Title: US-09-596-101a-3

Sequence: 1 MRMTLVVLCILFEGSEFAFTD.....DDEGDNNAQVDYIRVYKRN 384

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Maximum number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_unclassified:*
13: SP_vertebrate:*
14: SP_virus:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2240	100.0	384	5 077072	077072 Eisenia fo
2	716.5	32.0	361	5 09064	09064 pacifastacu
3	713	31.8	489	5 02660	02660 strongyloce
4	703	31.4	395	5 017492	017492 anopheles g
5	390.5	17.4	495	5 09N189	09N189 bombyx mori
6	369.5	16.5	877	5 045095	045095 bacillus ci
7	357	15.9	492	5 09VVR5	09VVR5 drosophila
8	353	15.8	1324	2 059328	059328 clostridium
9	350	15.6	276	2 052754	052754 rhodothermu
10	350	15.6	435	2 068641	068641 oerskovia x
11	329.5	14.7	487	5 09N198	09N198 manduca sex
12	323	14.4	494	5 09N198	09N198 drosophila
13	314	14.0	467	5 017233	017233 bombyx mori
14	307	13.7	306	2 051333	051333 oerskovia x
15	302	13.5	481	5 096363	096363 hyphanta
16	295	13.2	490	5 09VSR4	09VSR4 drosophila
17	295	13.2	490	5 09N189	09N189 drosophila
18	287	12.8	646	2 060039	060039 thermotoga
19	286.5	12.8	422	2 092690	092690 flavobacter

20	284	12.7	642	2 09W1	09W1 thermotoga
21	275.5	12.3	285	2 09F340	09F340 streptomyces
22	257.5	11.5	297	1 073951	073951 pyrococcus
23	249.5	11.1	410	5 09VVR4	09VVR4 drosophila
24	249.5	11.1	461	5 09N189	09N189 drosophila
25	239	10.7	422	2 09W1	09W1 streptomyces
26	209	9.3	673	5 027082	027082 tachypleus
27	190.5	8.5	478	2 09L185	09L185 streptomyces
28	184.5	8.2	845	2 09W1	09W1 clostridium
29	179.5	8.0	466	2 088021	088021 streptomyces
30	148	6.6	364	14 084415	084415 parametium
31	138.5	6.2	294	2 007242	007242 mycobacteri
32	132	5.9	419	3 09WS3	09WS3 schizosacch
33	125.5	5.6	1144	1 09V0D4	09V0D4 bacillus su
34	125	5.6	857	2 031530	031530 pyrococcus
35	121.5	5.4	424	3 060019	060019 pacifia rho
36	116.5	5.2	878	2 09KER4	09KER4 bacillus ha
37	115	5.1	284	10 P93670	P93670 hordeum vul
38	111	5.0	290	2 09MOD2	09MOD2 arabidopsis
39	110.5	4.9	290	2 09R119	09R119 aeromonas s
40	110.5	4.9	856	3 074170	074170 aspergillus
41	108	4.8	282	10 080803	080803 arabidopsis
42	108	4.8	636	3 013789	013789 schizosacch
43	108	4.8	720	3 006472	006472 saccharomyc
44	107.5	4.8	545	2 084907	084907 zebellia ga
45	107.5	4.8	725	3 09P670	09P670 neurospora

ALIGNMENTS

```

RESULT 1
077072 PRELIMINARY; PRT; 384 AA.
ID 077072;
AC 077072;
DT 01-NOV-1998 (TREMBLER, 08, Created)
DI 01-NOV-1998 (TREMBLER, 08, Last sequence update)
DE 01-NOV-1998 (TREMBLER, 08, Last annotation update)
DE COELOMIC CYTOLYTIC FACTOR 1.
GN CCF1.
OS Eisenia foetida (Common branding worm) (Common dung-worm).
OC Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC Lumbricina; Lumbricidae; Eisenia.
CX NCBI_TaxID=6396;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98406152; PubMed=9733802;
RA Beschin A., Biles L., Gomez J., De Baetselier P., Timmermans M.,
RA Reyers H., Brys L., Gomez J., De Baetselier P., Timmermans M.,
RT "Identification and cloning of a glucan- and lipopolysaccharide-
RT binding protein from Eisenia foetida earthworm involved in the
RT activation of prophenoloxidase cascade."
RL J. Biol. Chem. 273:24948-24954(1998).
DR EMBL: AF030028; AAC35887.1;
SQ SEQUENCE 384 AA; 44322 MW; C90B5C94003BAD6D CRC64;

```

Query Match 100.0%; Score 2240; DB 5; Length 384;
Best Local Similarity 100.0%; Pred. No. 1e-182;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

0Y 1 MRMTLVVLCILFEGSEFAFTDMDQYIHWODEFDYDGAKKMOHEVATGSGNEFOLYTD 60
Db 1 MRMTLVVLCILFEGSEFAFTDMDQYIHWODEFDYDGAKKMOHEVATGSGNEFOLYTD 60
0Y 61 GANSEVRDGLKFIKPTLLADNINPOTGAPFGTFMYNGVLDMWAMYGACTNDNNGCYRT 120
Db 61 GANSEVRDGLKFIKPTLLADNINPOTGAPFGTFMYNGVLDMWAMYGACTNDNNGCYRT 120
0Y 121 GAAGDIPPMASARVTFQKYSFTHGVRVYHAKMPYGDWLMPIAIVMLPEDWVYGGWPRSGE 180
Db 121 GAAGDIPPMASARVTFQKYSFTHGVRVYHAKMPYGDWLMPIAIVMLPEDWVYGGWPRSGE 180

```


[illegible]

RN	[1]	SEQUENCE FROM N.A.
RP		STRAIN-KINSHU X SHOWA; TISSUE=HEMOCYTES;
RC		MEDLINE=20138243; PubMed=10671539;
RA		Ochiai M., Ashida M.;
RT		"A pattern-recognition protein for beta-1,3-Glucan. The binding domain and the cDNA cloning of beta-1,3-glucan recognition protein from the silkworm, bombyx mori.";
RL	J. Biol. Chem.	275:4995-5002(2000).
DR	EMBL; AB026441;	BAA92243.1; -.
KW	SIGNAL.	
FT	CHAIN	17 495 BETA-1,3-GLUCAN RECOGNITION PROTEIN.
SO	SEQUENCE	495 AA; 55802 MW; 09243039F77456BF2 CXC64;
Query Match	17.4%;	Score 390.5; DB 5; Length 495;
Best Local Similarity	30.6%;	Pred. No. 3e-25;
Matches 114;	Conservative 57;	Mismatches 122; Indels 79; Gaps 19;
OY	26	IWDQDEDD--YFDGAKWOHEVATGGNSFEQLYTQDANSFVRDGKLFKEFTLLADNIN 83
Dd	185	LLEDFQFNIPHGRGIWPEVKFPGEPPFENVLYSD--NAEVNDGKLIIKPATLE--- 238
OY	84	PQTGAPEPTIDMYNVLDVMAMYGACTWTDDNG--CYRFGAAD--IPPMSARVTFQKY 140
Dd	239	---SKYEDEDVROS-LD--LSEKCTVTGTAOCLRE-ASGPLILPITIAKISTRHOE 289
OY	141	SETHGRRVVVAHKMPGDMLPAIMWLPRDWYVGWP-RSGEDIILETTIGNRDKNTGCEF 199
Dd	290	AFKRGVRVIRAKMRPGMDLPBELLPELRDNITYGVANNYSAGLIKIASVGNMEF----- 342
OY	200	LGIQMGSTMHGCP---GWMDNRYYLTSLPKHSDDMWNYGDNFHTFWFPMWSFNGLRFYVD 256
Dd	343	-----SKLIYAGRPIMTSSDPYRSFYELKENIGYESNN--NDFHNNTLEWRPDGITLLVDG 394
OY	257	EN-----QALLDPYRYLIDANPMWVPWFWEKGKFWLPQYENDNMAGSTNLAPDQNFH 309
Dd	395	ESYGEIKRGESEFPYN-----ANSYKVA-----AQO-----WLGITAPDELFFY 435
OY	310	FILNAVAVGTGTFPIDGCINGRDPALQRPSNGDMYNDAMRKFPDANGMMKMTWDDGCD 369
Dd	436	VSIGLVAVAGIREFSDD-----ISKMKPKNS--ARKAMLKFPDARSQWPFPTMEDD-- 482
OY	370	NNAMQVDYIRAY 381
Dd	483	-SALQVDYKVVF 493
RESULT	6	
ID	045095	PRELIMINARY; PRT; 877 AA.
AC	045095	053369;
DT	01-NOV-1996	(TREMBLrel. 01, Created)
RT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)
DE	BETA-1,3-GLUCANASE BGH PRECURSOR.	
GN	BGH.	
OS	Bacillus circulans.	
OC	Bacteria, Firmicutes; Bacillus/Clostridium group;	
NCBI_TaxID=1397;	Bacillus/Staphylococcus group; Bacillus.	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN-IAMI165;	
RX	MEDLINE=94033858; PubMed=7764221;	
RA	Yamamoto M., Aono R., Horikoshi K.;	
RT	"Structure of the 87-kDa beta-1,3-galactanase gene of Bacillus circulans IAM165 and properties of the enzyme accumulated in the periplasm of Escherichia coli carrying the gene.";	
RL	Biosci. Biotechnol. Biochem.	57:1518-1525(1993).
DR	EMBL; D17519; BAA00469.1; -.	
SO	SEQUENCE	495 AA; 55802 MW; 09243039F77456BF2 CXC64;

DE ENDO-1,3(4)-BETA-GLUCANASE (EC 3.2.1.6) (ENDO-1,4-BETA-GLUCANASE)
DE ENDO-1,3-BETA-GLUCANASE) (LAMINARINASE).
GN LICA.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DSM 1237;
RA Schwarz W.H., Schimming S., Fuchs K.P., Staudenbauer W.L.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOLYTIC ACTIVITY OF 1,3- OR 1,4-LINKAGES IN
CC BETA-D-GLUCANS WHEN THE GLUCOSE RESIDUE WHOSE REDUCING GROUP IS
CC INVOLVED IN THE LINKAGE TO BE HYDROLYSED IS ITSELF SUBSTITUTED AT
CC C-3.
DR EMBL; X89732; CAA61884.1; -
DR HSSP; P14090; IULO.
DR InterPro: IPR001119; -
DR InterPro: IPR003305; -
DR Pfam; PF00395; SLH; 3.
DR Pfam; PF02018; CBD; 6; 4.
DR Hydrolase; Glycosidase.
SQ SEQUENCE 1324 AA; 148202 MW; 038E40F9E09538AB CRC64;

Query Match 15.8%; Score 353; DB 2; Length 1324;
Best Local Similarity 27.2%; Pred. No. 1.6e-21;
Matches 100; Conservative 51; Mismatches 93; Indels 124; Gaps 16;

QY 22 DQYIIVMODED--YFDGAKQHEVTANGGNSFQLYTQDGSANFVDRGKLFIPYLLA 79
DB 425 EEMRLVMSDEFGSEIINMANMSYDPTNGRMNGEVSQYTON--NAVIXGALVIEAR--- 479
QY 80 DNINPQGAAPGTGTFMYGVLDMVAMYGACTNTDNNGCYRTGAADIPRPMASARYTFOK 139
DB 480 ---KEDITEPGEYTHY---SSKLTITGK 503
QY 140 YSFTHGAVYVAKKPVGDMLPAILMPLPDM--VYGGWPRSGEIDITETIGNDFKNTGGE 198
DB 504 KSMYKGEIKAKKPKQGQGPIMPATIMPEDEPFYGTWPKCGSIDIMELLGHP----- 556
QY 199 FLGIQKSGTMMKPGMDNDNRWLTSLPKHSDMNYGNFTFTWDSNGILRFYVDEN 258
DB 557 ---DKYGTGTHFGEPRKESQGYT-LP---EGQTFADDFHYSTIEMERGEIRMYIDGR- 607
QY 259 QALLDVPRPLDANPWWVDFFWEMGKPMPLROYENDNPMAG--TNLAPPDONHFLTNVAV 316
DB 608 --LVHV-----ANDW-----ISRDPYLAADITYPAFPDQNFLLTNISV 644
QY 317 GGTNGFTPDGCIINGDPAQKPMNSGDMYNDAMRKFFDARGNMKWTWDEEDNNAMQYD 376
DB 645 G-----GWWG-----YPDETTVF-----PQGMVVD 665
QY 377 YIRYKRN 384
DB 666 YVRYOKD 673

RESULT 9
ID 052754 PRELIMINARY; PRT; 276 AA.
AC 052754;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
DE LAMINARINASE.
GN LAMR.
OS Rhodothermus marinus.
OC Bacteria; CFB group; Rhodothermus group; Rhodothermus.
OX NCBI_TaxID=29549;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN-IT1278;
RA Boriss R., Krah M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF047003; AAC69707.1; -
DR InterPro: IPR000757; -
DR Pfam; PF00722; Glyco_hydro.16; 1.
DR PROSITE; PS01034; GLYCOSYL_HYDROL.F16; 1.
SQ SEQUENCE 276 AA; 32028 MW; 6D303D5FB04B48DE CRC64;

Query Match 15.6%; Score 350; DB 2; Length 276;
Best Local Similarity 28.5%; Pred. No. 4e-22;
Matches 116; Conservative 39; Mismatches 92; Indels 160; Gaps 19;

QY 1 MRETLVILC-LFEGEFAFTWDO-----YHIVODEDYF--DGAKQHEVTATGG 50
DB 2 MGVAFILCSLF--GCSILDGDPYRLPHWELVMSDEFGNGLPBAKWDYGVGHWG 59
QY 51 NSEFQLYTQDGSANFVDRGKLFIPYLLADNINPQGAAPGTGTFMYGVLDMVAMYGACT 110
DB 60 NOELQYTR-----ARIEARVG-----GGVLIIEARRESYE 91
QY 111 NFDNNGCYRTGAADIPRPMASARYTFOKYSFTHGAVYVAKKPVGDMLPAILMPLPDM 170
DB 92 GREYT-----SARLYTRGKASWTYKGFETIARLPSSRGTPAILMPLPDRQ 136
QY 171 VYGG--WPRSGEIDITETIG-NRD-----FKNTGGEFLGIQKSGTMMHSGMDNRYW 221
DB 137 TGSAYWPNPGEIDIMEHGFNDVYHGTAKYHHLGTOGGS----- 182
QY 222 LLSLPKHSDDMNYGNFTFTWDSNGILRFYVDE-----NQALLDVPRPLDANPWW 275
DB 183 -IRVPARD-----FHYATIEWPTEIRMEFVDDSLYRFPNERL-----TNP-E 225
QY 276 VDFWEMGKFMPLROYENDNPMAGSTNLAPPDONHFLTNVAVGDTNGTIPDGCINRGDPA 335
DB 226 ADMRW-----PDPQPHLLMNTLVAGSTWG-----QQGVDP- 257
QY 336 LQPMNSGDMYNDAMRKFFDARGNMKWTWDEEDNNAMQYDTRYK 382
DB 258 -----EAFPAQ-----LVYDYRYVR 273

RESULT 10
ID 068641 PRELIMINARY; PRT; 435 AA.
AC 068641;
DT 01-AUG-1998 (Tremblrel. 07, Created)
DT 01-AUG-1998 (Tremblrel. 07, last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)
DE BETA-1,3-GLUCANASE II.
GN BGLII.
OS Oerskovia xanthineolytica.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Micrococciaceae; Cellulomonadaceae; Cellulomonas.
OX NCBI_TaxID=1826;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LL G109;
RL Enzyme Microb. Technol. 13:71-75(1991).
RA Ventom A.M., Asenjo J.A.;
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LL G109;
RX MEDLINE=96409238; PubMed=8814220;
RA Parrado J., Escudero P.R., Conejero-Lara F., Kotik M., Ponting C.P.,
RT Asenjo J.A., Dobson C.M.;
RT "Molecular characterisation of a thermoactive beta-1,3-glucanase from
RT Oerskovia xanthineolytica";
RL Blochim. Biophys. Acta 1296:145-151(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-LL G109;

```

RX MEDLINE-96250169; PubMed-8659924;
RA Ferrer P., Hedegaard L., Halkier T., Diers I., Savva D., Asenjo J.A.;
RT "Molecular cloning of a lytic beta-1,3-glucanase gene from Oerskovia
RT xanthiolytica L6109. A beta-1,3-glucanase able to selectively
RT permeabilize the yeast cell wall."
RT Ann. N. Y. Acad. Sci. 782:555-566(1996).
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=LL G109;
RA Ferrer P., Andrews B.A., Asenjo J.A., Hedegaard L., Diers I.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF052745; AAC38290.1;
DR HSSP; P23904; 1AUK.
DR InterPro; IPR000757;
DR InterPro; IPR000772;
DR Pfam; PF00552; Ricin_B lectin; 1.
DR Pfam; PF00722; Glyco_hydro_16; 1.
DR Prosite; PS01034; GLYCOSYL_HYDROL_F16; 1.
DR SMART; SM00458; RICIN; 1.
SEQUENCE 435 AA; 46097 MW; 00F087BE644C0F58 CRC64;

Query Match 15.6%; Score 350; DB 2; Length 435;
Best Local Similarity 28.5%; Pred. No. 7.1e-22;
Matches 103; Conservative 30; Mismatches 96; Indels 132; Gaps 12;

QY 26 IWQDEFDYEDG-----AKQHEVTATGGNSFQLYTODGANSFYVDKLFIRPTLLAD 80
DB 57 LMSDFEDGAGSAPNPVANNHETGAGMGNALQNTASRANSAL-DGQ----- 105
QY 81 NINPQTGAPGEGTDFMYNGVLDMVAMYGACTNTDNGCYRTGAAGDIPASARVRFQKY 140
DB 106 -----GNLVITA-----RREGDSY-----ISAKRTTGKY 131
QY 141 SFTHGRVYVHAAPVGDMLPAIMLPEDWVYGGPSPSGEIDITETGNDFKNTGEPFL 200
DB 132 QPQYGRIEARIQIPRGQIPAFWMLGSGSPFPWSSGEIDIMENYGFEPHVRHG--- 187
QY 201 GIGKMGSTMHMGPGMDNRWLTSLPKHSDWMYGDNFHFWFMPSPNGLRFFVDENQA 260
DB 188 -----YH-GPGYSGGG-GITGMYTOHPQGSFADTFHFPAVDMKPGELTWVD--GQQ 236
QY 261 LLDVPPYPLIDANPMWVDFWEMGKFWLPQYENDNPMAGTILAPDQNFHFLNVAVGSTN 320
DB 237 FHVETBASVGANM-----VFQPFELLNVAVGQW 268
QY 321 GFLPDGICINRGDPALQKPSNGDWINDAMRKFFDANGKMKWTWDEGDNNAMQVDYIRV 380
DB 269 PGYPDG-----TQQLP-----QOMKVDYVRV 289
QY 381 Y 381
DB 290 Y 290

RESULT 11
Q9NJ98 PRELIMINARY; PRT; 487 AA.
AC Q9NJ98;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE BETA-1,3-GLUCAN RECOGNITION PROTEIN.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_Taxid=7130;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FAT BODY;
RX MEDLINE-20179841; PubMed-10713054;
RA Ma C., Kanost M.R.;

```

```

RT "A beta-1,3-glucan Recognition protein from an insect, Manduca sexta,
RT Agglutinates Microorganisms and Activates the Phenoloxidase Cascade.";
RT J. Biol. Chem. 273:7505-7514(2000).
DR EMBL; AF177982; AAF44011.1;
SQ SEQUENCE 487 AA; 54567 MW; 3BA4869C5ADE934A CRC64;

Query Match 14.7%; Score 329.5; DB 5; Length 487;
Best Local Similarity 28.4%; Pred. No. 4.5e-20;
Matches 106; Conservative 55; Mismatches 135; Indels 77; Gaps 17;

QY 26 IWQDEFDYEDGAKQHEVTATGGNSFQLYTODGANSFYVDKLFIRPTLLADNIN 83
DB 173 MLEFDNKNPLADGRITPPELMPFGEPPYVFNWYKEDNLHVGNLVIRMPYLV---- 228
QY 84 PQGAPGCTDFMYNGVLDMVAMYGACTNTDNG--CYRTGA--AGDIPASARVRFQKYS 141
DB 229 ----TAGEDAIWR-TLD---LSDRCTGLTAQCKRPPSALIVPVTAKINTKRTFA 280
QY 142 FTTHGRVYVHAAPVGDMLPAIMLPEDWVYGGMR-----SGEIDITETGNDFKNT--G 196
DB 281 FKYGRVETSAKMPRGDWLVPLQLEPVKNKG--IRNVSGILRVACYKNTETIKTLVG 338
QY 197 GEFYGIQKMGSTMHMGPGMDNRWLTSLPKHSDWMYGDNFHFWFMPSPNGLRFFVVD 256
DB 339 GPIMS-----EAEPYRTALKEFISNEPTEHNTLEMSPDALITMAYD- 383
QY 257 ENQALLDVPYPLIDANPMWVDFWEMGKFWLPQYENDP-----WAGTNLAPDQNF 308
DB 384 -----GIVYGRVTA-----PAGFGYEANQONVEAARWYQSGNIAPFDIMF 425
QY 309 HFLINVAVGCTNGFLPDGICINRGDPALQKPSNGDWINDAMRKFFDARGMKMTWDEG 368
DB 426 YISLQMDVGVYHER-PDEALN-----KPKKN--TATKAMVNNARSONNPTWLE-- 472
QY 369 DNNAMQVDYIRYV 381
DB 473 SERALLVDYVRV 485

RESULT 12
Q9NH80 PRELIMINARY; PRT; 494 AA.
AC Q9NH80;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE GRAM-NEGATIVE BACTERIA BINDING PROTEIN 1.
GN GNBPI OR CG6895.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim Y.S., Ryu J.H., Han S.J., Choi K.H., Brey P.T., Lee W.J.;
RT "Molecular cloning and functional analysis of Drosophila melanogaster
RT Gram-negative bacteria binding protein family.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF228472; AAF33849.1;
DR FlyBase; FBgn0040323; GNBPI.
SQ SEQUENCE 494 AA; 55455 MW; 461DF8E363EA312B CRC64;

Query Match 14.4%; Score 323; DB 5; Length 494;
Best Local Similarity 26.2%; Pred. No. 1.6e-19;
Matches 98; Conservative 65; Mismatches 135; Indels 76; Gaps 17;

QY 26 IWQDEFDYEDGAKQHEV-TATGGNSFQLYTODGANSFYVDKLFIRPTLLADNINP 84
DB 175 LLEFETFDQINESLWIDHVRPLDSKDAEFVLY--DG-KAVYHDCNLVTEPLMSS----- 227

```

```

0Y 85 QTAPEFGDE-MYNGSVLDWAMMGACTNTOJ--NGCY-----RTGAAIDIPMASARTRF 137
Db 228 -----YRDLJIANSRJD---LSRCGTGTHARIKCEJLHSTGSPSPGIMPIYTPR1KST 279
0Y 138 OKSFTHGRVYVHAHAKMPGVMLMPALMMLP-EDWVYGGMPRSGEIDIEFTGNRDEFKNTG 196
Db 260 ETFAVOYRIEIRAKLPBGMDIVPLLLEPTEYTGSGTSGGOLRYALARGNSVLMPR 339
0Y 197 GEFLLIGKMGSTMHGPMWMDNRWMLSLPKHSDW-----NYGDNFTHFEDWSPLNG 249
Db 340 GKLYD-----GNSLYGP-----VLSIDAHREDLMLSKRISHFGDFHTYSLDSSNR 389
0Y 250 LRFVYDENQALLDVPYPLIDANPMWYDFEWEMGKRWLPJOTENDNPMAGSTNLAPFDONFH 309
Db 390 LLEFVYDGVYGEMLNGFLEIDENP-----RMVGGSPMAPFDEMF 429
0Y 310 FILVANVGCTNGEFLPDGCINRGDPALOKPWSNGDMYNDAMRKEFDARGMKW- -TWDE 367
Db 430 ISLGVSYGEGFDE-----VDHLRTATEYEKPWAT-----IIPRSCSSTRADOMLPTWKQ 480
0Y 368 GDNAMOVYDIRVY 381
Db 481 ---ALKIDYRVF 490

```

```

QY 86 TGRF-FGTFDFNYNGVLDVVMNMGACTMTDNNCGCRITGAADI--PRMNASAVRFFQKXYSF 143
Db 200 QHMGEFLDDSTISGTLN---LFSGCTSS-ADACTCKQASGADILPPVYSGRI-TSIGAF 254
QY 144 HGRVYVNAHKMVGWMLPMPALMPLREDWVYGGMP-RSGEIDIETIGN-----RDFKNT 195
Db 255 YGVEIYRKLKIDQGMVLPPELLLEPFLKTKYSGMNTASGVVITACRGAHELYSGPNDSNT 314
QY 196 GGEFLGIQKMGSTNHHMPGMD-DYRMYLTSLPKHSDDMNNGDNHFFWFMSPNGLEFFV 254
Db 315 -----VYGGPIYMDBERENFSLTKRRRDSITSMGDSHTYTSVQMTDFIALSV 362
QY 255 DDENQALLDVYPLIDANPMWVDMWEMGKPMPLPOYENDNPMAGSTNLAPFDQNFHLNV 314
Db 363 DGEEMARVEAPRDAL-----PAVCNAPRHLILQASQANAPRDFDHITLGV 408
QY 315 AVGTNGEIPDGCINRGCDPALQCPWNGMDYINDAMK-----FFDAGNKKWTWDEEGDN 370
Db 409 AAGGITTEF-RDGSITSGS---VTKPW-----RDSAKKASVHEFRHMSDFPRMSOP--- 455
QY 371 NAMQVDYIRV 380
Db 456 -SLIVDPVKV 464

```

```

RESULT 13
ID 017233 PRELIMINARY; PRT; 467 AA.
AC 017233.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE GRAM-NEGATIVE-BINDING PROTEIN PRECURSOR (GNBP).
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_TaxID=7091;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 42-53; 56-83 AND 88-100.
RT TISSUE=FAF BODY, AND HEMOLymph;
RX MEDLINE=96353914; PubMed=8755572.
RA Lee W.J., Lee J.D., Kravchenko V.V., Ulevitch R.J., Brey P.T.;
RT "Purification and molecular cloning of an inducible gram-negative
RT bactericidal protein from the silkworm, Bombyx mori.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:7888-7893(1996).
CC -1- FUNCTION: MAY BE INVOLVED IN THE INSECT-IMMUNE PROTEIN. IT HAS
CC STRONG AFFINITY TO THE CELL WALL OF GRAM-NEGATIVE BACTERIA.
CC -1- SUBCELLULAR LOCATION: SECRETED IN HEMOLymph.
CC -1- TISSUE SPECIFICITY: FAF BODY AND TO A LESSER DEGREE IN
CC CUTICULAR EPIDERMAL CELLS.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION STARTS 6 HOURS AFTER AN INJURY
CC OR MICROBIAL CHALLENGE.
CC -1- INDUCTION: BY BACTERIAL INFECTION.
CC -1- SIMILARITY: TO H.CUNEA GRAM-NEGATIVE-BINDING PROTEIN.
CC -1- SIMILARITY: WEAK, TO FAMILY 16 OF GLYCOSYL HYDROLASES.
DR EMBL; U38591; AAB40946.1; -.
KW Signal; Glycoprotein; Insect immunity.
FT SIGNAL 1
FT CHAIN 19 467 POTENTIAL.
FT CARBOHYD 119 467 GRAM-NEGATIVE-BINDING PROTEIN.
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 467 AA; 52217 MW; 780652F89046B77 CRC64;

Query Match 14.0%; Score 314; DB 5; Length 467;
Best Local Similarity 27.3%; Pred. No. 8.9e-19;
Matches 101; Conservative 58; Mismatches 143; Indels 68; Gaps

27 WQDEDFVFDGAKQ-HEVYATGGSNSEFQLYTDGANSFPRDGKLPFKPILLADINPQ 85
:::11:::11:::11:::11:::11:::11:::11:::11:::11:::11:::11:::11
147 IFEEDFSDENWVMDIEDIPIYHEDEYFVSYSORNNLTVSTADNLMH-----NAKIQ 199

```

RESULT	14			
ID	051333	PRELIMINARY;	PRT;	306 AA.
AC	051333;			
DT	01-NOV-1996 (TREMBLrel, 01, Created)			
DT	01-NOV-1996 (TREMBLrel, 01, Last sequence update)			
DT	01-MAY-2000 (TREMBLrel, 13, Last annotation update)			
DE	BEA1-1,3-GLUCANASE IIA.			
GN	BGLIIA.			
OS	Oerskovia xanthineolytica.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
CC	Actinomycetales; Micrococcales; Cellulomonadaceae; Cellulomonas.			
OX	NCHI_TaxID=1826;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=LD G109;			
RX	MEDLINE=96345651; PubMed=8755914;			
RA	Ferrer P., Halkier T., Hedegeard L., Savva D., Diers I., Asenjo J.A.,			
RT	"Nucleotide sequence of a beta-1,3-glucanase isoenzyme IIA gene of			
RT	Oerskovia xanthineolytica LD G109 (Cellulomonas cellulans) and initial			
RT	characterization of the recombinant enzyme expressed in Bacillus			
RT	subtilis.";			
RL	J. Bacteriol. 178:4751-4757(1996).			
DR	EMBL; U56935; AAC44371.1; "			
DR	HSSP; P23904; IAJK.			
DR	InterPro: IPR000757; "			
DR	Pfam; PF00722; Glyco_hydro_16; 1.			
DR	PROSITE; PS01034; GLYCOSYL_HYDROL_F16; 1.			
SO	SEQUENCE 306 AA; 32835 MW; CD8DB8C1A6F8DC04 CRC64;			
Query Match	13.7%;	Score 307;	DB 2;	Length 306;
Best Local Similarity	26.2%;	Pred. No. 2.1e-18;		
Matches 95; Conservative	36;	Mismatches 97;	Indels 134;	Gaps
26	IWMDEDFYFGAK-----WQEHVATGGGNSFQLYTQDGANSFYRDKLIFKPTLLAD	80		
:		:		:
68	LAMSEDFGAAGSAPNPVMWHETGAGGWAELQNTTSHVNSAL-DCQ-----	116		
81	NINPOTGAPGCTDEFTNGVLDVWAMYGACTYTDNNGCRTGAAGDIPPMASARYTQKY	144		
:		:		:
117	-----GNLVITAL-----QESDGST-----TSARLTQGNV	142		
141	SETHGRVVVHAKMPYGDILMPALIMWLPEDWYVGGPRGSEIDITETGNRDFKWTGEFL	200		
:		:		:
143	QPFGRLEARLQIRPGGSGINSALFMVWGANLDPTRPPISGEIDIMENGNAPHEVHG----	198		
201	GIQKSGSTMHWGPGWD-DNRYWLTSLPKHSDMWYGDNFHTFMFDWSPNGRLFFVDENQ	255		

```

Db 199 -----TVA-GGYSQSDN--GIMGTQHPQGSFADDFHTFGIDMTPEGTTLVLD--GQ 246
QY 260 ALLDVEYPLIDANPWWVDWFEMWCKPMLPOYENDNPAGTINLAPFDONFHEILNVAAGT 319
Db 247 EYHRTVTTADVAGANQM-----VEDDPFLILNVAAG- 277
QY 320 NGTIPDGCIRGGDPALOKPMSNGDMYNDAMRKFFDARGNKMTWDEGNNANQVDYIR 379
Db 278 -----QMPGNPDATPEP-----QOKKVDYVR 299
QY 380 VY 381
Db 300 VY 301

```

```

RESULT 15
096363 PRELIMINARY; PRT; 481 AA.
096363;

```

```

01-MAY-1999 (T-EMBLrel. 10, Created)
01-MAY-1999 (T-EMBLrel. 10, last sequence update)
01-MAY-2000 (T-EMBLrel. 13, last annotation update)
DE GRAM-NEGATIVE-BINDING PROTEIN (FRAGMENT).
GN GNBPI.
OS Hyphantria cunea (Fall webworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Noctuidae; Arctiidae; Hyphantria.
OX NCBI_TaxID=39466;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99035790; PubMed=9818384;
RA Shin S.W., Park S.-S., Park D.-S., Kim M.G., Kim S.C., Brey P.T.,
RA Park H.-Y.;
RT Isolation and characterization of immune-related genes from the fall
RT webworm, Hyphantria cunea, using PCR-based differential display and
RT subtractive cloning.
RL Insect Biochem. Mol. Biol. 28:827-837(1998).
CC -1- FUNCTION: MAY BE INVOLVED IN THE INSECT-IMMUNE PROTEIN. IT PRESENT
CC STRONG AFFINITY TO THE CELL WALL OF GRAM-NEGATIVE BACTERIA.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION STARTS A FEW MINUTES AFTER AN
CC INJURY OR MICROBIAL CHALLENGE THEN STAYS RELATIVELY HIGH FOR 12
CC OR 24 HOURS.
CC -1- INDUCTION: BY BACTERIAL INFECTION.
CC -1- SIMILARITY: TO H.CUNEA GRAM-NEGATIVE-BINDING PROTEIN.
CC -1- SIMILARITY: WEAK, TO FAMILY 16 OF GLYCOSYL HYDROLASES.
EMBL AF023916; AAD09290.1;
KW Glycoprotein; Insect Immunity.
FT CARBOHYD 110 110 N-LINKED (GLCNAC...)(POTENTIAL).
SQ SEQUENCE 481 AA; 53014 MW; 49B2AC8C406929F6 CRC64;

```

```

Query Match 13.5%; Score 302; DB 5; Length 481;
Best Local Similarity 28.6%; Pred. NO. 9.7e-18;
Matches 107; Conservative 51; Mismatches 146; Indels 70; Gaps 19;

```

```

QY 27 VMODEPFFGAKWQHE---VTATGGNSSEFOLYTQGANSPFADGKLFKPTLLADNI 82
Db 155 IFEENFTREDVWOIYIPVYSTEPEFVSIOHLSODPYVA-VTGGNLRITPKL----- 208
QY 83 NPOTGAPFGTD-FMYNGVLDMWAMYGACTNTDNNGCYRTG-AAGDIPPMASARYTPOKY 140
Db 209 --QORMGFTDSSISYSSINT--FSGCT-APAACMKDANGASILPPVSGRI-TSKAF 261
QY 141 SFTHGKRVVAAKAVGMLPAIMLPEDWYGCNP-RSGEIDITETGNR-----DF 192
Db 262 AFTYGTGVFAKAKLPQGMWIPPEILPEFLKRYGSHYSSGVIKIASANGNRELTSGYTDY 321
QY 193 KNT--GGEPLGIQKMSSTMGMGPMQNDNRMYLTLSPKHSDDWNTGDNFHTFMDSPNG 249

```

```

Db 322 SNKMIFGSPVNMNQ-CYDTLLESKASSNGROW-----GDDEFHEVYLWAPER 367
QY 250 LRFVYDDENQALLDVPYPLIDANPWWVDWFEMWCKPMLPOYENDNP---WAGSTINLAPEDQ 306
Db 368 ITLSVDGVEMARVEPTAGLSGR-----FPOTCSKLPRTFLAAGTKMAPEDD 414
QY 307 NEHFLNVAVGSTNGFTPDGCINRGDPALOKPMSNGDMYNDAMRKFFDARGNKMTWMD 366
Db 415 HFYTLTGVAAGSTTEF-PDGVQTSGRP---KPMWNTG--SKAMLFWEEDMSFRAITNQ 468
QY 367 EGDNNAMQVDYIRV 380
Db 469 P----QLLVQYKV 478

```

```

Search completed: October 11, 2001, 15:56:45
Job time: 303 sec

```